SYNTHESE LIBRARY / VOLUME 178

ANALYTICAL PHILOSOPHY IN COMPARATIVE PERSPECTIVE

Exploratory Essays in Current Theories and Classical Indian Theories of Meaning and Reference

Edited by
Bimal Krishna Matilal and Jaysankar Lal Shaw

D. REIDEL PUBLISHING COMPANY
DORDRECHT / BOSTON / LANCASTER

ANALYTICAL PHILOSOPHY IN COMPARATIVE PERSPECTIVE

SYNTHESE LIBRARY

STUDIES IN EPISTEMOLOGY,

LOGIC, METHODOLOGY, AND PHILOSOPHY OF SCIENCE

VOLUME 178

ANALYTICAL PHILOSOPHY IN COMPARATIVE PERSPECTIVE

Exploratory Essays in Current Theories and Classical Indian Theories of Meaning and Reference

Edited by

BIMAL KRISHNA MATILAL
All Souls College, Oxford

and

JAYSANKAR LAL SHAW
Victoria University of Wellington, New Zealand and University of Hawaii, Honolulu

D. REIDEL PUBLISHING COMPANY

A MEMBER OF THE KLUWER ACADEMIC PUBLISHERS GROUP

DORDRECHT / BOSTON / LANCASTER

Library of Congress Cataloging in Publication Data

CIP

Main entry under title:

Analytical philosophy in comparative perspective.

(Synthese library ; v. 178)
Includes index.
1. Analysis (Philosophy)-Addresses, essays, lectures. I. Matilal, Bimal Krishna, II. Shaw, J. L., 1939-
B808.5.A533 1984 149'.943 84-20259
ISBN-13: 978-94-009-6501-0 e-ISBN-13: 978-94-009-6499-0
DOI: 10.1007/978-94-009-6499-0

Published by D. Reidel Publishing Company,
P.O. Box 17, 3300 AA Dordrecht, Holland.

Sold and distributed in the U.S.A. and Canada
by Kluwer Academic Publishers,
190 Old Derby Street, Hingham, MA 02043, U.S.A.

In all other countries, sold and distributed
by Kluwer Academic Publishers Group,
P.O. Box 322, 3300 AH Dordrecht, Holland.

CONTENTS

We dedicate this volume to
the memory of

Professor Kalidas Bhattacharya

who died suddenly
when the volume was in preparation

PREFACE

We are grateful to the authors who wrote papers specially for this volume and kindly gave their permission for printing them together. None of these papers appeared anywhere before.

Our special thanks are due to the first six authors who kindly responded to our request and agreed to join this new venture which we are calling 'comparative perspective' in analytical philosophy. In the introductory essay certain salient points from each paper have been noted only to show how 'comparative perspective' may add to, and be integrated with, modern philosophical discussion in the analytic tradition. Needless to say, any mistake, possible mis-attribution or misrepresentation of the views of the original authors of the papers (appearing in the said introductory essay) is entirely the responsibility of the author of that essay. The author apologizes if there has been such unintentional misrepresentation and insists that the readers should depend upon the original papers themselves for their own understanding. For typographical problems it has not always been possible to use the symbols originally used by the authors, but care has been taken to use the proper substitute for each of them.

Bimal K. Matilal

ANALYTICAL PHILOSOPHY IN COMPARATIVE PERSPECTIVE: AN INTRODUCTION

1.

The aim of this volume is to extend the horizon of philosophical analysis as it is practiced today. If two different streams of philosophical ideas that originated and developed quite independently of each other are found to be grappling with the same or similar problems and trying to find answers to similar questions and puzzles, this fact is by itself interesting enough for further exploration. Both contemporary analytical philosophy and the classical Nyāya and Buddhist tradition of India seem to be interested in the problems of knowledge and perception, the varieties of meaning and reference, the theory of inference and, the issue of psychologism. We wish to bring together these two very different streams and present them side by side if only to note, in the final analysis, their differences and contrasts. For it is also philosophically important to ponder why very similar puzzles evoke different responses from different people.

We need to say very little to introduce a contemporary problem, such as that of proper names, varieties and vagaries of reference, syncategorematic words, and modalities. But when such problems are raised in the context of classical Indian tradition, one needs to devote a lot of time in explaining the contexts of their origin and in providing the background material to make them intelligible. Besides, modern studies of classical Indian philosophical ideas often suffer from two disadvantages. First, the use of Sanskrit terms in parentheses becomes indispensable to signal the fact that these terms have acquired rather technical senses whereas their suggested English equivalent may be neutral and non-committal. It is expected that the reader should make allowance for this flexibility. Second, recent exegetical writings on Indian classical thought are often done in isolation and hence it fails to reach the contemporary philosophic audience. The resulting loss affects both sides. A non-Sanskritist finds a serious modern exegesis of classical Nyāya doctrine opaque. And if the exegesis is devoid of the depth of scholarship, it becomes too shallow and hence too plain and trivial. In this way a modern writer on the classical doctrine loses the vital criticism of the modern philosophic mind and a modern

B. K. Matilal and J. L. Shaw (eds.), Analytical Philosophy in Comparative Perspective, 1–37.

philosopher loses the benefit of the philosophic insight of the classical Sanskrit writers. We would attempt to mend the situation. For we are convinced that this will open up a new area of contemporary philosophic research.

Over the past fifty years in the Western analytical tradition one major debate has been concerned with the problem of intensionality and related matters. If we connect it with the issues in philosophy of mind we can even call it the debate about intentionality. Professor G.E.M. Anscombe (as well as P.T. Geach) has also made the philological point which reveals that the two words 'intention' and 'intension' are even genetically connected. By "related matters" we mean for example the dispute between realism (about universals) and nominalism, the theory of meaning versus the theory of reference (as W.V. Quine once called it), the admission of concrete versus abstract entities in one's ontology as well as the question of metaphysical necessity and essence. The essays presented in this volume deal mostly with such issues, and eight of these essays present comparative viewpoint or perspective by discussing generally kindred issues in classical Indian philosophy.

It was Franz Brentano who proposed (in 1874) the concept of intentionality as a characteristic that distinguishes what is mental and psychological from that which is merely physical. One significant aspect of intentionality is that the verbs of propositional attitudes and related psychological verbs denote psychological activities that are "directed upon" objects but such objects need not exist (nor need they be _causally_ responsible) in order to be intended by such activities. As R. Chisholm has put it, "We can desire or think about horses that don't exist, but we can ride only on those that do." (_Realism and the Background of Phenomenology_, Free Press, Illinois, 1960, p. 4).

The psychological component of the Brentano thesis, according to which, 'intentionality' has been explained as 'the reference to something as an object' (in Brentano's own words, 1874, ibid.) and taken to be 'a distinguishing feature of all mental phenomena' for no physical phenomenon manifests anything similar, has been more or less uncontroversial. But the ontological component of this thesis, viz., the doctrine of 'intentional inexistence' as being a _mode of being_ that may be attributed to the 'objects' of our thought, desires, search, etc., has been a rather controversial issue (and Brentano himself abandoned it later on, although this did not stop the controversy). In a modified manner, this controversy still survives in modern philosophical literature under a different guise. In fact, it is tied to a deeper philosophical issue that distinguishes realism from non-realism, physicalism from mentalism. (See also Lycan below.) More generally, the mind-body problem has been a live issue since the birth of

philosophical thinking.

I have already noted elsewhere (Matilal, 1968) that the psychological part of the Brentano thesis was well recognized by the Navya-Naiyāyikas of India. Thus Gadādhara (17th century A.D.) remarked that the 'self-transcending reference to some OBJECT' (sa-viṣayakatva) is the necessary mark of our psychological attitudes such as 'cognition' (which covers episodic belief and all other cognitive attitudes), 'desire/will' (icchā), and 'effort or intention to act' (kṛti). The Naiyāyikas being, however, hard-headed realists (as Brentano was), rejected the notion of a special mode of being to be attributed to the 'objects' of cognition or desire. But they showed their ambivalence. They regarded such OBJECTS as identical with either the real objects of the actual world or (where such objects do not exist) the composite objects which are constructed out of bits and pieces of the real objects and must therefore be analysed or broken down so that their segregated parts may be identified with such bits and pieces of reality.

Philosophers of both India and the West have tried to resolve the dubious ontological status of the 'objects' of psychological attitudes in various ways. Consider:

1. Tom is thinking about a unicorn.
2. Tom believes (i.e., misperceives) there to be a particular snake (when in reality there is a rope).

Despite the obvious differences here, I propose to treat 1 and 2 as similar and on a par as far as the question of the ontological status of 'the snake' and 'the unicorn' is concerned. The following four broadly possible answers are generally available:

A. An intentional SNAKE is produced there by the mental act itself. Hence there emerges an actual snake--distinct from the mental act--but it has a different mode of being---different from both being the actually existent snake and being the non-existent one (being nothing). This is generally compatible with Brentano's earlier view, as well as with the Advaita Vedānta view that such an OBJECT has a being which is inexplicable either as existent or as non-existent (cf. sad-asad-anīrvacanīya).

B. No object is produced by the thought or the mental act concerned. The so-called SNAKE or UNICORN in the context is either

non-distinct from the thought itself or through analysis and explanation identifyable with the 'furniture' of the actual world. This view seems to be compatible with the Nyāya _anyathākhyāti_, the Prābhākara _sat-khyāti_ and Brentano's later view of 'reism'.

C. The OBJECT is produced, i.e., thought-constructed, but not _actual_. It is distinct from the act which 'constructs' it, a _possible_ object, a _possible_ snake or a _possible_ unicorn, as witnessed by such later utterances as "It would have been a snake," or "There might have been unicorns in our world." This seems to be compatible with the views of many modern philosophers.

D. No OBJECT is produced, but an actual 'imaginary' object is involved. An 'imaginary' object means an _incomplete_ object, viz., an _incomplete_ SNAKE, which did not bite the person (for it could not have done so) but frightened him all the same. This could be close to the view of a Meinongian. Under some interpretation, this could even be closer to the Buddhist view of _ātmakhyāti_, or even a Nyāya realist may hold this position (see Chakrabarti below).

The upshot of all this is that it is extremely difficult, if not impossible, to resolve this ontological issue. An _incomplete_ SNAKE is imaginary for it cannot be matched by any actual snake. And such a SNAKE also violates the law of excluded middle, for we cannot decide whether it is, say, longer than two feet or not.

The matter can be approached in another way. We may ask: "How are we using the word 'snake' or 'unicorn' in 1 and 2?" It seems that we can say with more confidence how we are _not_ using them. For it is clear that we are _not_ using them _referentially_, as W.V. Quine has repeatedly emphasized. This would mean that these words are _not_ being used to designate or refer to an object of this _actual_ world. But can we answer positively? How are they being used? G. Frege's answer seems to have been that such terms in _ungerade_ contexts refer to what would be their senses (_Sinne_) in ordinary contexts.

Quine has insisted, on the other hand, that occurrences of such terms in these contexts are non-referential. He develops his notion of _referentially opaque_ context in which he wishes to contain all problems connected with not only

propositional attitudes (e.g., 'objects' of psychological verbs) but also modal operators. An opaque context is unlike the referentially transparent context, for it resists inferences through the substitutivity of identity as well as quantification from outside. E.g., from the truth

3. $\square$ (9 > 5)

we cannot derive the falsity

4. $\square$ (the number of planets > 5)

although 'the number of planets = 9' is true. This failure of substitutivity of identity is evidence, according to Quine, for the term being in a non-referential or 'not purely designative' position. Similarly, although 'Scott = the author of Ivanhoe' is true we cannot derive from the truth

5. Tom believes that Scott is the author of Waverly

the proposition

6. Tom believes that the author of Ivanhoe is the author of Waverly.

Failure of both quantification and substitution of co-designative terms indicates that these terms cannot refer to the ordinary objects of the world we know best. Quine's usual advice has been that we abandon quantified modal logic as well as quantified propositional attitudes or the logic of belief.

Faced with this challenge, modern writers have devised means to get around the problem in one way or other in order to develop quantified modal logics, logic of belief and epistemic logic. The modal logicians today introduce the notion of 'possible worlds' to explain the concept of necessity etc., and use the notions of 'rigid designator' and 'vivid designator' (in the belief-contexts) to avoid at times the problems of quantification and substitution.

The idea of necessity is explained as being "true in all possible worlds" (i.e. true under all state-descriptions, to use the old phrase of R. Carnap; true under all conceivable circumstances, to paraphrase nearly enough S. Kripke's interpretation of 'possible worlds'). A rigid designator is said to be one that designates the same object in any possible world where the object exists. If '_a_' stands for the term then

$(\exists\, \underline{x})\, \Box\, (\underline{x} = \underline{a})$.

Kripke has claimed that proper names of our language are rigid designators in this sense.

It seems that certain terms of our language (such as a proper name, say 'Scott') must be rigid designators in a certain sense, for otherwise we would not be able to make counterfactual assertions with the help of such terms, as we actually do. This is Kripke's intuitive argument in favour of the notion of rigid designators. A counterfactual such as "Scott might not have written Ivanhoe" would not mean what it means unless "Scott" designated the same person in both the actual world where he wrote Ivanhoe and the possible world where he did not. When such terms come in pairs such that they designate the same object, then even in modal contexts they lend themselves to the substitutivity of identity.

In the logic of belief or in epistemic logic, a vivid designator is the analogue of rigid designator (D. Kaplan, "Quantifying In", 1969). For the believer, Tom, a term is a vivid designator when there exists a specific thing that he believes (or knows) it designates. Such terms will therefore freely instantiate quantifications and are subject to the substitutivity of identity in belief contexts.

A Quinean extensionalist argues that semantic considerations of quantified modal logic invoke the notion of essence or essentialism. The concept of a rigid designator is sustained by the talk of 'possible worlds'. Quine continues:

> Talk of possible worlds is a graphic way of waging the essentialist philosophy, but it is only that; it is not an explication. Essence is needed to identify an object from one possible world to another.
>
> (Theories and Things, Cambridge, Mass., p. 118)

The extensionalist insists that both notions, that of rigid designator and vivid designator, are dependent upon some context or other, and empty otherwise. Our notion of necessity is context-dependent. For relatively only to a particular inquiry, some predicates may be treated as playing a more basic role than others, and these may be treated as necessarily so. It may be necessary for a mathematician to be rational but it is not necessary (essential) for him to have two legs intact. But if the mathematician is also a cyclist it is necessary for him to be two-legged. Only with regard to such background groupings and other information, we can distinguish between 'essences' and 'accidents'. Quine writes:

> For that matter, the very notion of necessity makes sense to me only relative to context. Typically it is applied to what is assumed in an inquiry, as against what has yet to transpire.
>
> (Ibid., p. 121)

Similarly the notion of vivid designator is context-dependent. Quine argues:

> The notion of knowing or believing who or what someone or something is, is utterly dependent on context. Sometimes, when we ask who someone is, we see the face and want the name, sometimes the reverse. Sometimes we want to know his role in the community. Of itself the notion is empty.
>
> (Ibid., p. 121)

Modal logicians may try to salvage the notion of 'essence' by appealing to natural science. A 'possible world' is, as Kripke emphasizes, not a product of science fiction, but only a vivid way of phrasing the counterfactuals. (Some however may not balk at even the inflated ontology of 'possible worlds'!) If the demand is one of clarity and intelligibility of the modal and intensional contexts, then the device of possible worlds seems to do its job admirably well. But the extensionalist argues that this is not enough, for the blown-off dust now settles upon 'rigid designator', 'contrary-to-fact possibilities', etc. Quine comments rather facetiously:

> Let us recall then that some of us have deemed our contrary-to-fact conditionals themselves wanting in clarity. It is partly in response to this discomfort that the current literature on possible worlds has emerged. It is amusing to imagine that some of us same philosophers may be so bewildered by this further concept that we come to welcome the old familiar contrary-to-fact conditionals as a clarification, and are content at last to acquiesce in them.
>
> (Ibid., p. 173)

Kripke's influential papers on necessity and the logic of proper names have given rise to what is called now the New Theory of Reference. This New Theory is in part a revival of the old Millian idea that proper names are non-connotative. The name is said to directly latch on to the object named

without the intervention of any description. In this sense it has been considered to be a refutation of the Fregean idea that proper names have sense (Sinn). But this component of the theory has been controverted already by Michael Dummett and others.

The New Theory is also vocal in rejecting the traditional theory of meaning based upon the extension-intension distinction. It is however difficult to pin down the sort of intensionalism that is repudiated. It is not presumably the kind of intensionalism that Carnap developed using his technical relation of logical equivalence as his criterion of identity for intensions (see his Meaning and Necessity, Chicago, 1956). Ordinarily, intensions are sometimes regarded as mental entities of some sort (let us recall our difficulties about determining the ontological status of the 'objects' of propositional attitudes). They may be regarded as concepts associated with the words used by the speakers, or indeed the concepts the speakers have "in their heads." The extensionalist like Quine may take some comfort in this repudiation of such mental entities by the New Theorists. But not quite. For the New Theorist endorses metaphysical necessity and essential traits (see also Boër below).

The notion of rigid designator invokes the following question: "What does it mean to say that an individual is the same in different possible worlds?" However this should be distinguished from asking for a criterion of transworld identification of the individual in question. For the latter is an epistemological question. It has often been thought that this question vitiates the modal logicians' strategy. But, as L. Linsky has argued, the epistemological question need not be answered at all in order to make sense of possible world semantics for modal logic (Names and Descriptions, Chicago, 1980, p. 144). The first, the 'metaphysical' question, Linsky says, has a trivial answer. For indeed, to claim that an individual of one possible (actual) world is identical with another individual of another possible world is no more (or less) mysterious than claiming that they both share certain essential traits (whatever such traits may be) and differ in having many other changeable properties. In actual world we do identify Scott (for instance) under different circumstances at different times (or, as an Indian philosopher would add, in his childhood, youth and old age) without being actually able to articulate the 'individual essence' of Scott, which would presumably be a clear criterion for such re-identification.

Kripke has given strong arguments to show that identity statements such as "Tully = Cicero" are, if true, necessarily true. The apparent oddity of this claim can be easily removed. For Kripke insists that they are neither analytic nor a priori, but simply a posteriori necessary propositions. The notion of necessity in this context is metaphysical but

the notion of a priori or a posteriori is epistemological. Due to similar considerations "Water = H_2O" is claimed to be necessarily true. Under any counterfactual situation (or in 'Twin Earth') water will be the same, H_2O, and not some liquid XYZ that might have all the _superficial_ properties that we associate with actual water. The point is that if our science determines that water is H_2O or that gold is the metal whose atomic number is 79, then nothing counts as a possible world in which water isn't H_2O or gold is not that metal whose atomic number is 79. We can talk thus about metaphysical necessity and essential attributes.

Indeed our Kripkean modal logician accepts 'essences' of some natural kinds such as gold and water. The extensionalist may not find the point entirely unacceptable as long as such essences are identified with natural traits of some chemical or physical elements. Besides, our extensionalist is not strictly a nominalist, he is not, so to say, an 'ostrich nominalist'. He accepts classes as real, for one can give extensional criterion for identifying classes. He can espouse reality of certain universals (or 'essences' if you like), for they may be an integral part of the physical theory that uses them--the same physical theory that uses such items as atoms, electrons, stones and sticks. But he emphasizes nevertheless that one must try to state clearly the criterion for the individuation of these universals or essences in order to avoid the 'shadow' entities like the universals of the medievals. The adequate (intelligible) criterion, the extensionalist insists, should be that universals are identical if and only if their instances are the same. (Universals are therefore either _extensional_ class-properties or natural, physical traits, identifyable or observable.) Curiously enough, this is reminiscent of the _tulyatva_ 'sameness' criterion of two real universals in Nyāya as expounded by Udayana (in the eleventh century A.D.) while he was underlining the distinction between _real_ universals (_jāti_) and _bogus_ universals (_upādhi_). A real universal, according to Nyāya, is a simple, further unanalysable property shared by many. It is an 'unconditional' (natural) property of some natural kind or other. The extensionalist's demand coincides with another Nyāya intuition, according to which no real universal needs to be countenanced if the application of some general terms (e.g. "beast") can be explained in terms of some observable criterion or other. Beasthood (=_paśutva_) is thus analysed as being marked by dense body-hair and bushy tails etc. Notice that our extensionalist would reject all such universals as would lack the said extensional criterion. For where no extensional criterion can be given on demand the relevant attribute would, according to the extensionalist, be (in some sense) intensional and would cry out for clarification. "It is not clear then what kind of entities we are talking about

in this context," he would say.

Classes are admitted by the extensionalist for various reasons. One among them is making sense of taxonomy in a science like zoology. Classes, similarity notion and dispositional terms are, according to him, disreputable but practically indispensible. They are like "unredeemed notes" (W.V. Quine, "Natural Kinds", in S.P. Schwartz (ed.), Naming, Necessity and Natural Kinds, Cornell, Ithaca, 1977, p. 174). With the maturity of a branch of science, some of these turn respectable and in principle superfluous. It is however not absolutely clear why an unanalysed underlying property or similarity notion would be superfluous if and when a scientific theory would clear it up. If the notion of 'intelligence' is suddenly explained by the relevant branch of science in terms of proteins or colloids or whatever, why does it turn superfluous? Is it because the mystery about an intangible entity is all of a sudden gone? Or is it because it can no longer be a real universal in the traditional sense? It may be mentioned that this attitude is also shared by a Naiyāyika in India. Uddyotakara (6th century A.D.) used the example of pācaka ('a cook' or 'a chef') to argue against the Buddhist critique of real universals in order to make the point that not all such general terms or common nouns would deliver unanalysable real universals. For example, when we suddenly learn in our adulthood that chefs are not a special species of humans but only those who successfully went through a period of special training in the courses of culinary art, chef-hood become superficial, i.e. an upādhi, a bogus universal. The Nyāya contention that beasthood (paśutva) is not a real universal (jāti), as explained above, may also be similarly explained.

The extensionalist wants to restrict all talks of essences to natural traits of elements (admitted in our physical theory). But the non-extensionalist is not inclined to restrict our talk of essence in this way. The extensionalist assimilates the old nominalistic constraint by claiming that what is considered usually as an essence or an essential attribute of an object (apart from the natural and identifiable traits of object accepted by our physics) depends more or less upon the context or the discourse or the 'domain'. As we have already noted, it may be essential for a mathematician to be rational but if the same person is a bicyclist it would be essential for him to have two legs (or to be a 'featherless biped'). This intuition is also shared by a Naiyāyika. For he can use his notions of 'delimitor' (avacchedaka) and 'relevant distinguisher' (viśeṣaṇa) to underline the context-dependent character of such attributes. The Naiyāyika would say that while the occurrence of the property of being a mathematician in a person (the ādhāra or location) is delimited by the property of being rational and that of being a

bicyclist is delimited by bipedness, neither rationality nor bipedness would essentially characterize the person-as-such outside of some given context or other. He raises the question of real universal only when the term is either a natural kind term or a metaphysically fundamental term such as 'substance' and 'action'. Certain very similar artificial kind terms (such as 'pot' and 'cloth') are also admitted in the system as real universals but to avoid complication I would not enter into that question. A proponent of the New Theory of Reference however is inclined to tolerate such essences as cathood and many other such properties. He discusses the question of reference of gestalt terms, underlying trait terms and many similar terms and allows a mild form of essentialism (see S. Boër below).

The New Theory of names is also called a 'causal' theory. For one thing, the reference here is determined by something like a causal chain connected with a so-called original 'baptism' ceremony in a speech community and not by any associated description. The fix in this theory is not by the 'degree of fit' requirement, as some Description Theorists would like to have it, but by causal origin. A name is given to a person in a 'baptism' or a natural kind term ("water") is introduced by means of a paradigm. In this regard, we may be reminded of what the Naiyāyikas of India said: a term is introduced by a name-giver pointing to the object and saying "Let this word denote this object," where the intention of the name-giver (in some cases, God) becomes the śakti or 'power' infused so to say into the word such that it can denote the said object. This is what they call the 'conventional' (samketaja or sāmayika, cf. Vaiśeṣika-sūtra 7.2.24) fixing the referents of words.

The causal link with the original baptism is however much more complicated in most cases, as has been recognized by the proponents of the New Theory. But although some sort of causal link with the original referent is undeniable, it does not mean that this 'causal' theory is entirely free from intentional reference and mentalism. As G. Evans has argued, the relevant causal chain is not between the present use of the name and the original referent but between the body of knowledge that we associate with the name and the referent ("The Causal Theory of Names", in S.P. Schwartz, op. cit., pp. 213-234). In other words, the original referent should be in some way the source or causal origin of the body of information we associate with the name. This explanation, curiously enough, preserves the intuition of the Description Theorists and at the same time compatible with the modalist's doctrine that some identity statements (like "Tully = Cicero"), if true, are necessarily true. Some however would like to see the 'causal' theory interpreted as completely intension-free. But this would require much more than what is at the present

moment available in the literature.

The extensionalist concedes that the problem of intensional context with regard to the verbs of propositional attitude is not easily dispensible. For one wishes to maintain a logical difference between Tom's believing that there are spies and his suspecting somebody to be one of them, but the extensionalist's analysis in this regard has not been very successful. It is even suggested sometimes that for an extensionalist such distinctions may not exist (W.V. Quine, _Theories and Things_, p. 121). Quine concludes:

> We can end up rejecting _de re_ or quantified propositional attitudes generally, on a par with _de re_ quantified modal logic. Rejecting them, that is, except as idioms relativized to the context or situation at hand. We remain less cavalier toward propositional attutudes than toward modal logic only in the unquantified or _de dicto_ case, when the attitudes are taken as dyadic relations between people or other animals and closed sentences.
>
> (Ibid., p. 122)

I shall submit for consideration the following comment on the dispute between the extensionalists and the modal logicians. We may dismiss "Scott could not have been Scott" as a possibility (as its being true in any possible world) for such a dismissal is guaranteed by the logical falsity of "$\underline{a} \neq \underline{a}$". But can we in the same way dismiss "Scott could not have been human" as a genuine possibility? (For its analogue "$\underline{a}$ ≠ an $\underline{F}$" could hardly be a _logical_ falsity.) I submit that there is another strong intuition which can entertain the latter possibility: in a possible world, Scott was a cat. If we reject this as being impossible we are then back to the notion of individual essences being context-dependent. If a proper name designates (in the manner of a _logically_ proper name _a la_ Russell) an 'uncoloured' particular, an 'unclothed' individual, then we have to assume (with Quine) that if the individual identified by such a name, is an 'uncoloured' object, it remains an open question whether such predicates as "is a man" (or "is a cat") are true of it or not. Quine seems to understand the situation in a similar manner (note his manoeuvre to treat all names as predicates and put the burden of reference upon bound variables only), and hence he is right in emphasizing the _emptiness_ of the concept individualhood (the individual essence of an _unclothed_ individual). In this case, the essential or necessary trait of being Scott would depend upon our decision as to whether to consider the identified individual as a particular human being or not. However, once we

identify an individual by a name as a particular man, he becomes a (particular) contingent object of this actual world, and hence surely we can talk about his essential or necessary properties.

To put it simply: if an individual is identified (through, say, being the value of a bound variable) and we are not told yet whether he is a man or a balloon or a cat, we can very well say that its individual essence is 'empty' unless it is 'contextualized'. But mostly the names of our language identify their objects in a more specific mannner (as belonging to some 'category', as a man, as a cat or as a building), and in such cases it does not seem unnatural to talk about the necessary or essential properties of such a particular. An essential property of Scott is thus being the man he was, being born of the parents he had (or being born of those particular ovum and sperm). In short, a particular individual must have some essential traits, but if we know him only as an individual captured only as a value of a bound variable, its individual essence would certainly be empty. Names usually identify specific particulars, not the 'naked' ones. Hence talk of essence in such contexts may not be empty.

Quine's disavowal of the notion of 'possible worlds' has not also been very successful. He simply says that he is content with the old contrary-to-fact conditionals. This hardly amounts to a proper argument for refutation. In semantics, on the other hand, talk of 'possible worlds', missed chances, construction of hypothetical situations, alternative futures, potential dangers, etc. does seem to offer a sort of intuitive understanding. However the advantage of this device may be overemphasized and one may be eager to extend the border of semantics to include ontology. And here the warnings of the extensionalists may be well justified. In brief, if certain intuitively grasped distinctions are rejected as being only mentally induced, then the extensionalist's program for a successful 'flight' from essences and intensions may be complete. But the obvious question is: At what price? Care should be taken to ensure that we are not throwing away the baby with the bath water.

2.

2.1 Intension and Metalanguage

Recent researches have shown how one can resolve the problems with the verbs of propositional attitude (as well as modality) by either using intensional models or resorting to metalinguistic predicates. Some logicians prefer the latter. But the following paper of Cresswell shows that there is in fact not a great deal of choice that we can make between the two viewpoints, for they are intertranslatable. It is only a

matter of individual preference of the philosopher concerned whether to follow the intensional vocabulary or the metalinguistic vocabulary. For there is no great philosophical advantage in one which is not available in the other.

Cresswell is one of the well-known exponents of modal logic and possible worlds semantics for natural languages. In the present paper he argues that not only the necessity operator in modal systems can be interpreted as a metalinguistic predicate of sentences (Bryan Skyrms, 1978) but also any intensional operator can be construed in this way. Pigs do not fly in our world. But there are obviously two ways of making "pigs fly" true. We may either change the facts and make pigs fly or use the word "pigs" as we use "birds" (and change the meaning of "pigs" thereby). In favour of the intensional language Cresswell claims that such intuitively obvious distinctions (viz., the difference in extension caused by a change in facts and the difference in extension caused by a change in meaning) are better handled in an intensional rather than an extensional language. Extensionalists such as Quine would deny that there is any such difference here that needs to be accounted for. But this may be, as I have already noted, of a piece with Quine's program for the flight from intension.

A move to change the meaning of "pigs" is however a move to a different intensional model. This is different from expressing the problem by change of worlds in the same model. Cresswell considers how it is possible to express a class of intensional models by a family of classes of metalinguistic models. He also argues that there is at least one way in which an explicitly metalinguistic semantics and a two-dimensional semantics for propostional attitudes can be shown to be equivalent.

The problem lies however in our interpretation of the phrase "change of meaning". How are we to understand it? It is true that what "pigs" mean in our world (the world we know best) imposes no constraints on what it could mean in another world. This sort of Humpty-Dumpty-ism is presupposed in the intensional models incorporating 'possible world' semantics. But in making the sentence "pigs fly" true in another world we cast 'sidelong glances' at how we use the words "pigs" and "fly" in <u>this</u> world and not in that other world. This is why, Cresswell suggests, R. Stalnaker regards his works on intensional language as pragmatics rather than semantics.

The result of Cresswell's paper is that there does not exist a great deal of philosophical difference between the two 'viewpoints', the intensional and the metalinguistic, prevalent in the interpretation of modalities, propositional attitudes, etc. One may add that both therefore form a common front against the extensionalists who, like Quine, recommend excision of all talks about intensional differences which

cannot be captured properly by an extensional language.

2.2 The 'Syncategorematic' Way

Gochet has very succinctly described the modern controversy over the notion of predicate and quantification. He generally approves of the Quinean approach to predicates, which he rightly calls the syncategorematic treatment of predicates. However he points out that it would be misleading to claim that general terms or predicates are fully syncategorematic like the 'iota' sign in Russell's theory of Description, for they have extensions, or in R.M. Martin's language, they are related to individuals by the semantic relation of 'multiple denotation'. Gochet chooses a special class of predicates, "∈", "⊂" and "=", for which he offers, more appropriately, a syncategorematic account. He calls them copulae. This will also include N. Goodman's mereological predicates such as "is a part of". These copulae predicates do not designate, they are part and parcel of the machinery of predication. This falls well within the line suggested by Quine's remark in Word and Object that fundamental forms of predicates and other predicates may be treated as verb-like expressions. Gochet argues that the copulae are only 'conceptual or linguistic duplications'. This account fits well with the prevalent nominalistic intuition which rejects generally the objectivity of universals or concepts.

Of many important issues raised in this context, we may select only two which are relevant to the general theme of this volume. Firstly, Gochet rejects G. Küng's notion of a concrete property as well as the related notion of 'inherence'. These notions have apparently been offered as constituting an alternative answer to such question "Why is a certain predicate sign true for some individuals and not for others?"--an answer that could presumably be compatible with the Quinean distrust for any 'abstract' non-extensional entities, i.e., the objective universals of the medievals. A reasonable account of the notion of a 'concrete property' may not however be absolutely difficult. Even without following the line of Küng and without using the notion of qualitative identity (in Küng's sense), one may still say with Nyāya that the adjectives in our language ascribe particular instances of a so-called property to the thing which they qualify. A particular instance of a property need not be a 'mysteriously' concrete property, but simply (in some sense of the term 'concrete') a concrete feature or a PARTICULAR whose presence or residence in another presumably concrete object may then be spoken of. For example, Nyāya interprets colour terms such as "red" as expressing residence of a particular instance of red (comparable to a particular red-tint in a picture) in an object (a ball) that is said to be red. It is not

mysterious to say that the predication in "the ground is seeded" means that the ground has seeds in it and seeds are as concrete as they come. Why call them properties if they are not abstract? This seems to beg the question. To avoid confusion we may call them concrete 'features' or following Nyāya let us call them objective locatees (cf. dharma or ādheya). The common characteristic between the concrete locatees and abstract properties (such as cowhood) is that they are both occurrent (vrttimān) and hence locatees. The connecting link between a locatee and its location need not always be 'inherence'. If by 'inherence' we mean what is called samavāya in Nyāya, a naturally inseparable relation between a locatee and its location (substratum), then such a connection exists only between certain groups of locatees and their substrata. According to Nyāya, it exists between the abstract property cowhood and a cow, between a particular colour-instance and the coloured substance, and between a particular action and the acting object. Such a link is posited where the two entities connected are deemed as ontologically distinct but inseparably tied together from the origin. (Notice that cowhood can be manifested in nothing else but a cow, and the particular colour-tone and the particular act cannot originate without the substrata [the things] in which they actually originate and 'belong'.) There may be other connectors such as the one of physical contact between the seeds and the ground in the given example. Each physical contact is treated here as a particular, and hence another connector ('inherence') can very well be spoken of as introducing the required tie or copulative. Or, sometimes the connection may simply be a subjective connection constructed for the purpose of some interest-oriented description or a heuristic device (e.g. ownership relation or svatva-svāmitva in Nyāya). Or it may be simply a conceptual duplication in the way Gochet describes the copulae. In the last case the locatee may be a subjective entity or a trivial and artificially constructed entity which is ontologically non-distinct from the location. For example, from such sentences as "Cowhood is a universal" we may construe a trivial locatee-attribute, universalhood, which is obviously ontologically non-distinct from the substratum, cowhood. Nyāya admits fundamental universals (natural kinds) as real, but not universalhood (as real). Here we may talk about a copula predicate or, what comes to the same thing, a svarūpa-sambandha in Nyāya, which would combine the substratum with the superficial locatee, universalhood. This will then have a syncategorematic treatment, for it is declared as an integral part of the mechanism of predication. The notion of conceptual or linguistic duplication helps also to resolve the so-called paradox of relation (which is usually known as Bradley's paradox and which was explicitly formulated in the Indian tradition by Dharmakīrti

in 7th century A.D.). The copula tie stops the infinite regress before it begins.

This brings us to the second point. One may not accept a syncategorematic account of ordinary predicates or adjectives for we may construe them as attributing properties or instances of properties to things, but even so, some sort of copula or syncategorematic predication becomes necessary for any language. Just as we cannot have well-formed sentences in many languages without having (overtly or covertly) verbs in them, we cannot have a logic of predication without admitting some verb-like or syncategorematic predicates. Goodman's "is a part of" is such a predicate. So is the Naiyāyika's "is qualified by". These copulae predicates, in Naiyāyika's language, svarūpa-sambandha 'self-same connector', do not name properties or components of individuals; they are only linguistic duplication. And such duplication seems unavoidable in any system that we try to construct.

2.3 Naming and Believing

Lycan connects the so-called paradox of naming (whether proper names have senses or connotations) with Kripke's puzzle about belief and suggests a two-scheme hypothesis to resolve the dilemma between Millianism and Russellianism. Since Saul Kripke's most influential arguments against the so-called Description Theory of Names which Lycan calls Russellianism, a decade of writers developed different kinds of semantical theories to tackle the problems of proper names, referential opacity and related matters. What seems rather interesting from a comparative viewpoint is that Lycan favours a theory of mental episodes which he construes as brain episodes, to which he assigns 'intentions' or the sorts of structure that we associate with bits of public language. Although this is a sort of physicalism which is in consonance with the view of most modern philosophers, there is nevertheless a distant resonance here of a Naiyāyika's account of a cognitive episode having a sort of structure verbalizable in bits of public language. Lycan's use of dot-quotes can be comfortably set against the Naiyāyika's structural description of the cognitive episodes in terms of what is called 'object-hoods' (viṣayatās). However, in this way of looking at matters, one should be wary of what is called 'tacit' belief, and hence Lycan rightly restricts his discussion to what he calls occurrent or episodic belief. In this respect too Lycan is closer to Nyāya where jñānas are always treated as episodes.

Lycan develops a two-way scheme. One component here is the computational scheme of J.A. Fodor where the mental tokens of a person S are individuated on the basis of their inferential and computational roles as well as their behaviour-causing roles. The other component is derived from some

acceptable semantical theory where the mental tokens are individuated, say, according to the sameness of truth-conditions. In this way Lycan suggests a resolution of, or an answer to, the Kripke-type 'hard' question: If a person utters _sincerely_ "Cicero was bald but Tully was not," and behaves accordingly, does he irrationally believe a contradiction?

We may hazard a guess about an alternative way that an exponent of Nyāya may prefer. We should of course note that 'believe' is not exactly a part of the Nyāya vocabulary, but we may supplant the episodic beliefs here by the Naiyāyika's awareness-episodes or cognitive events. If this is done then we may proceed in this way. If the above token or utterance (of S) is a true verbal representation of the structure of S's awareness-episode at that particular moment (confirmed, let us say, by the sincerity of tone and the ostensive behaviour) then we have to say that S is not cognitively aware of a contradiction here, for S's mental [Cicero] and [Tully] play different 'computational' roles even though both [Cicero] and [Tully] are representations of one and the same person. Nyāya would talk about Cicero-objecthood and Tully-objecthood as two distinct objecthoods. (We may represent them by square brackets: [Cicero] and [Tully].) According to Nyāya each objecthood has a two-way determination. It is determined on the one hand by the object itself and on the other by the unique cognitive episode (in fact by its structural peculiarity) to which it owes its origin. What is called the 'computational role' seems very similar to what is determined by the structural peculiarity of an episode. Hence it is possible for two such objecthoods to be distinct from each other although they may be, unknown to S, grounded in, i.e., determined by, the same object of this world. If we presuppose, as we must in the context, that S is rational at that moment, we are not allowed to infer from the given occasion that S irrationally believes a contradiction (whatever status we may ascribe to such beliefs). The second determination from the side of the structural peculiarity of the episode would account for this distinction between the Cicero-objecthood and the Tully-objecthood. Strictly speaking, the Naiyāyikas would use the notion of 'delimitor' (_avacchedaka_) to underline the distinction: It would be said that although the purported substratum or _dharmin_ (location) of both baldness and the lack of it is one and the same ontological entity, the episode presents its epistemic counterpart, the 'role' or 'mode' of the substratum as _delimited_ by two different objecthoods, the Cicero-objecthood in one case and the Tully-objecthood in the other. (See below.)

Lycan however favours a sort of representationalism. The believer's mental "Tully" and mental "Cicero" are, he says, brain-state "representations" of (caused by?) the same physical object--representations playing different computational

roles. Nyāya does not seem to endorse such a representationalistic view. For the 'objecthoods' are not separate occurrent realities. They are heuristic abstractions from the context of a cognitive episode. They may be substitutes for intentionalities. They are indeed queer sorts of entities, as Gadādhara noted in his Viṣayatāvāda (see also my The Navya-Nyāya Doctrine of Negation, 1968, chapter 2). They are more like grammatical entities, their role playing is always determined by the given construction. The same stick, for example, becomes an agent, an instrument or a location depending on the construction: "The stick touches the ground," "The man touches the stick," and "My hand rests on the stick." Lycan's concern is to vindicate physicalism (or a sort of functionalism, perhaps) against the constraints of intentionality. This however still remains an open question.

2.4 Reference and the New Theory

Boër reflects on various aspects and implications of the New Theory of Reference initiated by Kripke, Putnam and others. He concentrates upon the use of mass nouns and the count nouns as substance and kind words in order to vindicate the theory against the intensionalist's objections. He discusses what W. Goosens has called underlying trait terms and notes that substance-names such as "water" and "gold" and natural kind terms such as "cat" are usually said to conform to this New Theory. Arguably the category of such terms can be extended to include some artifact terms such as "coffeepot". If we introduce the notion of gestalt-terms, we can handle the problems of common names for plants and animals, and we need not claim that they have to correspond always to the taxonomical definitions of Final Science. Ordinary kind-names of plants and animals may be deemed as parasitical on the related gestalt-terms and on a certain amount of antecedent interest-relative sorting of items within the gestalt. Most artifact-terms are however hard nuts to crack. Boër argues that in artifacts the underlying trait or nature can be given by a reference to their specific functions (Kornblith). Another suggestion of Boër's is that we treat some of these terms as hybrid terms in whose definitions some underlying trait term or other would be present.

Although the New Theory is deemed well-suited to handle theoretical terms, there has been some misgiving about the success of Putnam's program in explaining the semantical difference between such terms as "phlogiston" and "oxygen"--or in explaining the behaviour of 'electron' (where a shift in reference has occurred in the history of the concept). To dissipate these and other criticisms voiced by Fine and Enc, Boër points out that we need not have difficulties in introducing high-level theoretical terms via descriptions employing

antecedently introduced underlying trait terms and various laws constituting them. Just as one can exploit the notion of hybrid-terms and gestalt-terms in order to avoid intensionalism and appeal to synonymy, one may follow a similar procedure here.

Underlying trait terms are claimed to be rigid designators in subject position. This claim, it is argued, is best understood as saying that each such term (qua singular term) refers to a property having an instance in the actual world and to the same property in every possible world in which it has instances. The extension of such a term (qua predicate) at a world w is that of the corresponding property at ẇ. Boër argues that while the New Theory yields a weak form of essentialism about kind-identity, it entails no strong form of essentialism about kind-membership.

It is however not absolutely clear how our property-talk in the context would not entail the strong essentialism. If the implication is that we need not distinguish between necessary and contingent properties unless in some sort of interest-relative way, then this will be a welcome move in favour of extensionalism: No property is essential to the object unless we take it in the background of some theory or other. This will then be compatible with two other claims of the New Theory. 1) Scientists are sometimes discovering necessary truths--not just contingencies. 2) A scientific investigation into the atomic, chemical or biological structures of some kinds of things is an investigation into the so-called essence of that kind.

Moravcsik has defended the notion of intension of general terms against the New Theorists. I shall comment on only two points raised in this connection. Moravcsik emphasizes the crucial role of 'intension' in our learning, and gathering competence to use, the relevant mass and count terms. Several Indian philosophers have argued for a long time that unless we refer to the 'quality' (guṇa) or jāti or 'class property' by such common nouns we would be guilty of 'losing count' and 'promiscuity' (ānantya and vyabhicāra). The simple point is that if a term is used to pick out many such individuals with which its (presumably conventional) connection (marriage?) has not yet been established, it would be 'promiscuous' in its application or usage. We have to explain the obvious: How even a child, once he has been introduced to the use of the term with reference to a particular elephant, would cry out "An elephant!" when a new elephant is presented to him? (See K. Bhattacharya below.) It is difficult to disassociate the role of 'intension' from our teaching and learning method. The New Theorists use the notions of paradigm and similarity and appeal to the notion of the Final Science for associating properties with terms. Roughly the reference is to be determined by causal links in the chain of events leading back to

the original referent. But if G. Evans is right, the relevant causal chain is between the body of knowledge we associate with the names and the referents. If the same technique is extended to general terms, will the theory be completely intension-free or intention-free? It may be right to claim with Boër that in introducing an underlying trait term like "gold" we leave it to the future scientists to identify the underlying trait in terms of which the metal's (gold's) individuation and persistence are to be understood. But the intensionalist need not deny this point as he can claim with Moravcsik that knowledge of individuation and persistence are part of intensions. It is connected with what it is to know and understand a language.

Our appeal to 'intensions' is often scorned, for they are usually thought to be in the 'heads' of the speakers or language-users. Or, it is argued that talk of intensions asks us to assume something intractable beyond the usually observable and readily inferable phenomena. But our intension-talk need not imply such shadowy entities in all cases. The New Theorists' talk of essences and properties may be seen as only an explication of old intensions of some well-known general terms. Our nagging suspicion about intensions may be partially dissipated if we think of them as heuristic devices or even as partly objectified posits until some scientific theory is forthcoming to make them full-fledged objects (properties). We may invest them with a half-way objectivity through the general consideration that they are not totally private or intrasubjective but more often than not intersubjectively accessible. A similar half-way objectivity is attached to Frege's Gedanke. Besides, a satisfactory analysis of the verbs of propositional attitude has been the source of constant worries for an extensionalist. The general point is: it is very difficult to get rid of the intensional idioms entirely.

Another objection to the New Theory considers the use of underlying trait terms in belief contexts. One may say that "John believes that water is a fluid" can differ in truth-value from "John believes that H_2O is a fluid," and this is to be explained in terms of some semantic difference between "water" and "H_2O" (and such difference is assumed to be difference in intensions). Boër argues that if we extend Kripke's line ("A Puzzle about Belief") to underlying trait terms, we can show that the foregoing "substitution failure" objection is based upon an inconsistent set of assumptions, hence cannot be used to support intensionalism.

There may be other puzzles about belief where the intensionalists could have an edge over the New Theorists. Suppose John has never taken his first lesson in Chemistry or Science while James has. John utters (sincerely, etc.), "Water is a fluid." But James utters, "H_2O is a fluid." Although John

and James believe in the *same* fact or state of affairs (a point on which the New Theorist and the Intensionalist may agree) they *represent* this object *differently* to themselves, and "intensions" may be invoked to explain this difference in the way it is represented or conceived.

Let us see how a Naiyāyika would view such matters. He would use, as already indicated before, his notion of delimitors or determinants (*avacchedaka*) of the application of the two words, "water" and "H_2O" used in the said utterances. To wit: although the referents of the two terms are the same, the Naiyāyika claims that the delimitors of the *role* of their being referring words or their being 'powerful' to designate something (cf. *śakyatā*) are different. In one case it is presumably the 'body of knowledge' which the user associates with the term "water" and in the other the 'body of knowledge' the other user associates with "H_2O". Certainly these two bodies of information cannot be identical although they may overlap in the given case. Thus the question still remains: Have we been able to avoid reference to 'intensions' altogether? Besides, the intensionalist can argue that the notion of *proper* translation as opposed to interpretation cannot depend simply on pure references, it should preserve intension in order to meet the standards of proper translation. Otherwise the distinction between translation and interpretation will vanish.

2.5 Belief and Logical Certainty

Jackson provides a recipe for the solution of one 'easy' version of the 'surprise' examination paradox. The elegant manner in which it is presented requires no comment. We may direct attention to what is said here about a general feature of belief, justified or reasonable belief. The idea is that we may be justified in believing that *p* while acknowledging at the same time that should certain things turn up differently later on we would not be justified in believing that *p*. If this means that I am justified in believing that I am eating an apple now for here is the apple I am holding in my right palm and taking bites but if this turns out to be a perfect dream later on I should not be so justified, then it contains the ingredients of well-known sceptical questions. In other words, one has to assume that one's justified belief at any given moment may turn out to be false should an unknown factor be revealed later on. This defeasibility feature of justifed belief plays a crucial role in solving what Jackson calls the easy examination paradox. But there is a hard version of the paradox which cannot arguably be solved in this way. For the hard version uses the notion of *certainty* in place of the notion of justified belief. Presumably in this version, at every stage prior to the surprise exam it is certain that

there will be an exam and that it will be a 'surprise' exam. How are we to understand this notion of certainty? Jackson says that it arguably lacks the above-mentioned defeasibility feature.

Certainty in this version is not however the usual psychological certainty. It is an 'ideal' or (if you like) 'logical' certainty. Such a certainty-episode seems to be equivalent to a knowledge-episode, according to at least the Nyāya notion of a knowledge-episode. For otherwise this certainty cannot lack the said defeasibility feature. To deny this point would amount to rejection by one stroke of the well-entrenched tradition of sceptical arguments against the possibility of knowledge. I may, for example, be (psychologically) certain absolutely that I am eating an apple now but still I may be intimidated by the sceptical possibility of Descartes' devil or Hilary Putnam's mad scientists who manipulate our 'brains in a vat' or Vasubandhu's idea of all our life-experience being simply a superdream. A non-defeasibile certainty tends indeed to be non-subjective. K. Bhattacharya would probably say (see his article below) that such certainties are being understood purely in terms of the given!

Jackson says that in fact the paradox can be solved even if there is certainty that there will be an exam but it cannot be solved in that way if there is in addition another certainty that it will be a surprise exam in the required manner. Now the certainty about the surprise element seems to be a different kettle of fish.

Certainty about the 'surprise' element may sometimes be read as the subject's certainty (knowledge?) about the subject's ignorance about the actual time (or day) of exam prior to such time. How do we understand the phrase "the subject's ignorance"? Ignorance is in fact an intractable property of the subject. The examinees may simply be certain that they do not now know the actual time or day, or they may be certain from the beginning that their ignorance is ineliminable. If the first, they would try to find out by arguing as they actually do in any version of the paradox. But since they presumably reach the conclusion (a certainty-episode) that occurrence of the exam along with the 'surprise' element is impossible, their two previous certainties are destroyed thereby and thus they re-introduce their own 'ignorance' or dubiety (for certainties are destroyed)! Then the surprise exam can be given even on the last day for it will be a surprise for those girls who argue in the above manner (kith and kin of Quine's solution?). If however the examinees accept the fact that they would never know the time prior to it they would not even undertake to expose a putative contradiction in the examiner's statement, and then the exam can be given at any time and still be a surprise!

2.6 Action and Intention

Segerberg presents an axiomatized model for actions with a description of its formal semantics to explicate what he calls a primitive notion of action within the broad tradition of intensional logic. Action cannot be always understood in terms of change interpreted as pairs of prior and posterior states. A richer notion of action should incorporate the agent's intention or a set of intentions attributable to the agent. Segerberg singles out an operational intention from the intention set and connects it with what he calls a routine. The agent runs such a routine to carry out the intention. The agent's skimming through the members of the intention set is called a 'deliberation walk'. A completely successful deliberation walk issues in an operative intention, the realization of which will realize all the intentions in the intention set.

Armed roughly with such conceptual apparatus Segerberg deliberates over the question raised by A. Goldman with regard to the concept of action. It is the problem of individuation of an act when presumably several actions happen to be performed by the same agent simultaneously. Examples are numerous. When I walk across the field, I touch the blades of the grass, disturb the flies, come near the river (my destination), generate movement in the air and tire my legs. Is this a set of several individual acts or only one act with several, in fact six, different descriptions? Goldman opts for the former while Davidson and Anscombe support the latter. Segerberg claims that his analysis combines features some of which support the many-action view while it can still preserve the intuition of the one-action view ("I did just one thing, acted under one operative intention bringing about a unique total change").

Some general observations on the issue may be in order. If we view the agent's action in terms of some physical change in the actual state of affairs, it is difficult to avoid the feeling that the one-action view is more correct. There may be several (correct) descriptions of the same physical event (brought about by an agent) just as there are always several possible descriptions of the same physical object, say a flagpole (a wooden object, a pointer, a work of art, a historical relic and so on). The dispute about whether I performed several other acts while I walked across the field cannot be easily resolved if we demand intentions be always behaviouristically interpreted. If however we countenance an ontology of mental occurrences, the situation may be quite different.

Goldman recognizes the intuition of the one-action theorists by construing the set of all (relevant) acts on a given occasion into a single act-tree. He says that the so-called single act, perhaps, corresponds 'to whatever "underlies" the

acts on a single act-tree' (p. 37). He uses the notions of basic act and level-generation to produce the act-trees where the basic act-tokens are at the bottom. Basic acts are mostly illustrated by bodily acts, moving of hands or legs. The role of intention is duly emphasized here, for nonbasic act-tokens are said to be connected with some basic act-tokens, and each basic act-token is connected with want-and-belief causation. Let me note that a Naiyāyika would find most of Goldman's claims for his theory of human action rather compatible with the Nyāya analysis of action. The notion of an "occurrent" want (a mental event) combined with occurrent beliefs explains the arising (causation) of a particular act-token. Nyāya likewise gives a causal analysis of human action in terms of three successive mental events: (1) an awareness-event (comparable to Goldman's occurrent belief) about x as a means to the agent's desired goal (iṣṭasādhanatā-jñāna, i.e., finding x attractive, nice, etc.) as well as thinking about x as acheivable or 'accomplishable' (kṛtisādhyatā-jñāna; "I can do x"), (2) a cikīrṣā or an occurrent want to do x, and (3) a kṛti or an emergent mental readiness for x. Here (1) leads to (2) and (2) leads to (3), which, in turn, leads to the particular act-token. A mental readiness to act is not a mental act (an act in Nyāya belongs to a different catgegory, kriyā as opposed to guṇa), or simply a volition. It is an emergent psychological property generated by the occurrent want to do x. Some Naiyāyikas suggest that (1) should include in addition another awareness of x as not being greatly harmful to the agent, in order to explain the lack of action on the part of the hungry person to eat poisonous food. In Goldman's account there are two elements jointly causing the act while in the Nyāya analysis there are three elements which are also causally related through sequence. (The Prābhākaras have argued however against the Naiyāyikas that in (1) an awareness of x as attractive, etc. is not necessary, for certain ritualistic acts are undertaken by the agent even when such acts are not attractive or subservient to any desired goal.) There is however agreement here (between Goldman and Nyāya) in the general theory of mental occurrences causing bodily act-tokens. Such agreement includes even the contrast of a 'logical connection' between the concept of a want and the concept of action and a 'causal connection' between particular wants and particular act-tokens.

I wish to make a further comment on the notion of change. Segerberg calls it an 'imperfect' notion of change which is defined as pairs of prior and posterior states. However, as Davidson has pointed out, under stricter notion of act-individuation there may be discovered many intermediate stages between this pair. A theory of human action is generally understood in the context of a broader notion of action identified either as change of states of physical bodies or

movements. Paradoxes with regard to the notion of movement are well-known (see Nāgārjuna in the Indian context). However if we atomicize action or motion to reach the infinitesimal the notion of action may finally vanish as it did completely in the Buddhist theory of universal fluctuations or momentariness: no body can move because it is always a new body at a new spatial location, and there are no prior-posterior stages of some single object but there are only prior and posterior stages. In this way actions and movements are all conceptual construction of a collection of sequential states. In Nyāya however actions are real particulars (act-tokens) although they are atomic and instantaneous. 'A body moves or acts' means in this theory that at each moment a new act-token is locatable in the body; it is generated leading (sequentially) to separation, dislocation and re-location (kriyāto vibhāgah, vibhāgāt pūrvasamyoga-dhvamsa uttara-samyogotpattih). An ordinary action over time such as cooking or walking means a conceptually constructed series of such moments during each of which an act-token is locatable in the body.

2.7 Meaning, Subjectivity and Objectivity

Bhattacharya (K) begins with some general problems connected with the notion of meaning. His concern however is to see how they are related to the traditional issues of philosophy. He argues that the relation between word and meaning, symbol and the symbolized, is neither fully objective nor fully subjective but we are required to attribute partial objectivity to such (presumably) collectively admitted and authoritatively handed down conventions. (The point is somewhat similar to Frege's contention about the distinctness of sense from purely subjective ideas on the one hand and physical objects on the other.) We do objectify even a fiction (or a fictional name) out of deference to the original name-user (as when we say, "But a Sherlock Holmes could have solved the case").

Towards the end Bhattacharya (K) makes an interesting observation about the so-called distinction between Western and Indian philosophical trends. He says that historically the West has been concerned with the speaker's meaning while the Indian philosophers have traditionally been concerned with the hearer's meaning. He generalizes them as two different fundamental attitudes in philosophy. One is characterized by an implicit dependence upon "I" while the other upon "this" and "he". The idea is that the former inevitably incorporates subjectivity into the notion of meaning while the latter need not do so. Apart from destroying the mistaken popular belief about the subjectivity and speculative nature of Indian philosophy, this point also explains why Indian (systematic) philosophers of language (and I would include even the Buddhist,

contrary to Bhattacharya's comment about Buddhism being an exception) have been uniformly "realists" with only a few solitary exceptions. For this reason one may note that such basic universals as 'elephanthood' (presumably the meaning of "elephant") are considered in Nyāya to be wholly objective and even observable in the particulars instantiating them. Bhattacharya (K) argues that even a child when he sees the second elephant in his life after being introduced with the first speaks out loudly "Elephant". The determinant (avacchedaka) of the 'power' of the word to denote is elephanthood (wherever it might fall in some later ontological analysis). This ties up with the dominant issues of this volume: universals, natural kinds and essence. This is also why Bhattacharya (S) would like to see such universals compared with essences (see below).

Western philosophers, Bhattacharya (K) says, paid over the ages so much attention to "I-thinking" or the speaker's attitude as dominating over the hearer's attitude, and hence in their recent theories of meaning they ought to have advocated the usual holistic attitude and preferred transcendentalism to empiricism. But for the last four hundred years there has been a volte face situation in the West (a "return of the repressed", perhaps?), for it is characterised by an aggressive (in Bhattacharya's description) moving away from the imaginative and transcendental philosophy. He calls it an aggressive hearer's attitude, and he further speculates that the old emphasis on the speaker's attitude as well as on the so-called individual freedom is still visible in this new trend. This does not coincide, therefore, with the old Indian passive acceptance and understanding of the GIVEN, of the Nature. Rather, it is an active suppression of the freedom of thought (lest it becomes unfounded speculation or wild imagination)--a freedom that paradoxically binds itself to Nature, the given, the matter. In classical India the Naiyāyikas came close to this attitude--for them knowledge is simply a happening, an episode rather than an active manipulation. The most typical of this "aggressive" philosophy is to be found, according to Bhattacharya (K), in the attempts of some philosophers as J.S. Mill who understood logic ("the functionality of thinking or speaking") purely in terms of the given, as only some highest generalisations of the given. Bhattacharya (K) notes that although the Millian view is no longer in vogue a modified form of this type of empiricism still dominates the analytical philosophy of today. We may add that this also accounts for frequent references to the Nyāya view in this volume while we deal with the views of modern analytical philosophy.

2.8 Universals

Bhattacharya (S) tries to sort out our perennial puzzles about universals and answer a few well-known objections about abstraction. His observations particularly on Quine and Wittgenstein are significant. What he develops as his Theory A seems to be based substantially upon the Navya-Nyāya view although there is no explicit reference to Nyāya except in the title of the paper. He suggests that the traditional difficulties with the concept of universal could to a large extent be avoided if we formulated the problem in a different way. Following the Nyāya view, for example, we may formulate the point in this way: Let universals be postulated, roughly speaking, as the 'reason or ground or basis for the application of general terms' to different individuals; this reason (pravṛttinimitta) is sometimes called 'meaning' (artha) and sometimes 'determinant of the meaning relation' (cf. śakyatāvacchedaka, more literally, delimitor of the denotative power of the word) in Nyāya. As I have already noted, this denotative power is nothing but a conventional 'power' infused into the word by some original "dubbing" situation. At some points Bhattacharya (S) seems to endorse the "essence" view of the universals, and in this way he comes closer to the view of some proponents of the New Theory of Reference.

Bhattacharya (S) seems to stop short of endorsing the Nyāya view completely. He rightly talks about a relation which should be part of the Nyāya postulation of universals as real entities over and above the particulars. (One may contrast this point about a real relation with that of Küng who has been criticized by Gochet.) However Bhattacharya (S) further notes that if universals are only posits or postulates one need not bother about asking such questions about the reality of relation between a universal and a particular. Explaining the example "marriage" (Strawson's example) he makes the point that neither all 'abstract' relations be universals nor all abstract expressions denote repeatable properties. There are so-called abstract expressions which in fact may stand for different unrepeatable features or facts. (Nyāya calls them either an upādhi or a svarūpasambandhaviśeṣa.)

While the concern of Bhattacharya (K) has been to see whether and in what way the modern worries about meaning in the West have been connected with the traditional issues of philosophy, Bhattacharya (S) has taken a particular theory of universals from traditional India (Nyāya) and reconstructed it in bare outlines in order to meet some age-old objections against universals. In fact a non-committal attitude towards the objectivity of universals may not entirely go against the spirit of some Naiyāyikas. As noted already, if the basis for application of a count name like 'chef' is exhausted by

referring to an objective feature (a particular) causally connected with the use of that term (e.g. training of each person in culinary art), we need not take chef-hood to be an objective universal. This is an old point made by Uddyotakara.

2.9 Psychologism

Mohanty deals with an important question, that of psychologism, as it may arise in connection with our discussion of Indian logical theories in contemporary terms. What he says here is particularly instructive and suitable for a volume such as this one. For when it is stated that Indian 'logicians' deal with not propositions but mental occurrences that we may call cognitive EVENTS, one usually faces the obvious question from those who are trained in the logical theories of Frege, Russell, etc.: Does it not reduce all Indian logical thinking to a sort of fruitless psychologism? Mohanty has shown that Nyāya supplies an alternative model for a logical theory that avoids the pitfalls of psychologism and extends the boundary of our logical thinking beyond the available models of the so-called Frege-Russell tradition.

Each cognitive episode or awareness-event has a structure or a form (which Mohanty is inclined to call 'intentional content' following the convenient terminology of Phenomenology). It is the same structure that Potter (see below) has called "contents" (and Potter has immediately emphasized that they are "always exhaustively composed of real items"). An awareness-event (as I usually call it) is both a particular and a momentary occurrent. They are much like the word-tokens or utterance-tokens and individuated by the moments of occurrence as well as by the subjects or persons in whom they arise or occur. But from the point of view of conceptual analysis and formulation of logical laws the identity condition of an awareness-event is given solely by its structural content. This structural content can be represented in language by an utterance, but it need not be so expressed in each case. However a structural description of such a content is always possible in language, and if the structural representations of two or more such events totally coincide (ignoring for the moment the problems presented by the indexicals), such events are treated as mere tokens for the 'same' awareness, i.e., the same type. In other words, for the purpose of logical and analytical study, the fact that they may arise at different times and in different persons is irrelevant. Mohanty successfully contrasts this notion with the Western notion of propositions and argues that this need not be confused with the 'propositions' of Western logic. (See also Potter.)

The so-called 'logical' laws of Indian philosophers, Mohanty says, are no doubt given in psychological terms, but since it follows an _eidetic_ psychology of cognitions, the

resulting theory of inference is not psychologistic in any pejorative sense. It is indifferently a logic and a psychology of inference. One of the consequences of such a psychology of inference and causal laws is that one has to say, as Nyāya does, that it is impossible to generate invalid inferences (from well-known logical laws). Mohanty argues that this is also a defensible position for the recent researches of Mary Henle show that the two blind alleys of psychologism and 'radical separation of logic from the study of thinking' should be avoided, if logic is to be brought closer to actual human thinking. In our study of Indian logical theories we can effectively steer clear of the two extremes: a totally depsychologized logic and a logical theory that is fully dependent upon probabilistic laws of psychology.

2.10 Speech-Acts and Jñāna

Potter takes up the issues where Mohanty leaves off. He argues that although the logicians, Navya-Nyāya in particular, talk about jñāna or a cognitive event (which he rightly interprets as an awareness-fact) it would not appear entirely queer to Western philosophers, provided we accept a speech-act model to understand it. The speech-act model was introduced by J.L. Austin in the Western tradition as an alternative to the propositional model of Frege and Russell. Today the main exponent of the speech-acts is J. Searle. With useful ingenuity Potter notes certain parallelisms between the speech-act model and the awareness-model of Nyāya. He then proceeds to suggest that as far as epistemology or logic is concerned with something like a speech-act model or an awareness-model it tends to be more natural than the "classical" propositional model. In support of this contention he cites the express uneasiness of certain outstanding exponents of the "classical" propositional model, regarding, for example, the exact notion of a proposition.

2.11 Mistranslation and Misunderstanding of nyāya in Modern Discussion

Daye approaches the problem of interpretation from a different point of view. He criticizes our usual practice of translating "anumāna" as 'inference' (as well as "pratyakṣa" as 'perception'), and argues that such a practice may mislead the non-Sanskritists into believing that anumāna (either svārtha or parārtha) is something derived from a properly formulated deductive schema. In his critique he tries to underline the difference between a deductive schema and an anumāna in the Buddhist (or even Nyāya) sense. His paper signals a timely warning for those among us, who may be overenthusiastic in discovering formal techniques in the Indian theory of

inference or <u>anumāna</u>.

Broadly speaking, in <u>svārthānumana</u> or what we sometimes translate as 'inference-for-one's-own-self' the Indian theorists exploit the psychological technique of drawing an inference or reaching a piece of knowledge based upon evidence, and in <u>parārthānumāna</u> or what we call 'inference-for-others' this technique is articulated in speech (language) so as to demonstrate its soundness to others. This 'psychological technique' however does not render the theory totally psychologistic in any pejorative sense (as has been argued already by Mohanty and Potter). It conceives of an ideal agent (or an ideal observer in the case of <u>pratyakṣa</u>) who obtains knowledge in this way from some sound evidence. The evidence will be infallible, it is claimed, if it fulfills the so-called <u>triple</u> condition (of the Buddhists) or the <u>quintuple</u> condition (of Nyāya).

Since the goal in a <u>pramāna</u> theory (epistemology?) is to state how we obtain knowledge (<u>pramā</u>) and not simply what constitutes a valid argument, this account of <u>anumāna</u> seems to be quite adequate for the purpose. However this provides a very different model for 'logic', if we may use the term at all in the context. Logicians who are interested only in formal validity or consistency lying in abstraction in a string of so-called propositions, a string that we call argument, need not be disappointed at this treatment of <u>anumāna</u> if they recognize the model here for what it is. It serves a different purpose in a different context (the <u>pramāna</u> theory). In other words, the theory of <u>anumāna</u> need not be an outlandish model for 'inference' even for a Westerner, provided we understand it as a part of what we may call the Indian program for systematic epistemology.

2.12 Designation and Related Designation

Siderits focuses upon the problem of how we derive the 'sentence-meaning' from the analysis of its components, words and their meanings. The two schools of Mīmāṃsā, the Prābhākara and the Bhāṭṭa, held opposite views on this issue. Briefly, the first contends that the sentence-meaning is given by the words directly, not through the <u>meanings</u> of the words concerned while the second believes that it is given only through the meanings of the words. Following Siderits, we can call the first the theory of 'related designation' and the second the theory of 'designated relation'. The nearest equivalent of this controversy in the West would be the modern discussion of the 'context principle' versus the 'composition principle' in semantics.

Siderits elaborates the Prābhākara view of 'related designation' which is based upon the strong intuition that word-meanings are seldom learned in isolation (i.e., apart

from the implicit or explicit context of a sentence). In fact the Prābhākara claims that they are learned in the context of some imperative sentence to begin with (e.g. "Bring a cow"). Hence whatever the words designate must always be overtly or covertly related to other things designated by other words which figure in the composition of the sentence. Mere juxtapostion of two or more words, if they designate only isolated meanings, cannot give rise to our awareness of a WHOLE, a connected fact, where components are interrelated, and blend together as one. Hence the Prābhākara argues that 'relatedness' cannot be severed from what a word designates. The opposite view maintains that word-meanings are primarily isolated objects of this world because they are learned from hearing the sentence through a process of analysis, isolation, comparison and correlation with such objects. The element of 'relatedness' in the sentence-meaning is given to us hearers by another means, a presumptive judgement based upon syntax, juxtaposition or structure (cf. <u>ākṣepa</u>).

The Prābhākara view is halfway between a holistic framework (e.g. that of Bhartṛhari) and an atomistic framework (of the Bhāṭṭas). It seems to claim that there may be two isolated objects in the world, bread and butter, but they cannot be the designata of the words, "bread" and "butter", in complete isolation. When we talk about them in our language we refer to them or the words designate them as <u>related</u> either to each other or to something else in the context. Siderits argues that the Prābhākara view enjoys the advantages of the Fregean 'context' principle without necessarily running into other difficulties which the Fregean principle supposedly engendered (such as positing of 'unsaturated objects').

<u>2.13 Empty Terms</u>

Chakrabarti deals with the familiar problem that is nicknamed (in the West) Plato's Beard. It is the old riddle of non-being. An important discussion on this riddle was the outcome of a controversy between the Nyāya and the Buddhist (logicians) over the issue of finding a 'negative' example negatively supporting a concomitance between 'universal' properties, i.e. properties that are present in all real objects. For the Buddhist such properties are momentariness and causal efficacy (which is by definition equivalent to <u>existence</u> in Buddhism), for they characterize everything that is real. According to Nyāya, such properties are knowability and nameability. But in spite of such difference in ontology, the logical form of the inference would be acceptable as equivalent in both systems: It is A because it is B, for all A's are B's (where A and B stand for bearers of universal properties in the sense already defined). This must be a logically impeccable inference. But the problem arises in the logical

theory when one is required to formulate the contraposed version, "All non-B's are non-A's," and support it by citing an example. An example of non-B must be a fictitious object. The dispute over this point raged over several centuries. It was at its peak with Udayana and Ratnakīrti.

Chakrabarti closely follows Udanyana's text to articulate the Nyāya position and uses Ratnakīrti's text occasionally for formulating the Buddhist view. At every step, he refers to the views of modern Western philosophers where they are relevant. This illuminates his textual analysis of Udayana--a method that could be used profitably for further analysis of other Sanskrit textual materials. He next selects Bhartṛhari and briefly deals with his (Bhartṛhari's) holism, as well as his idea that each substantival word would create its own object of reference. He talks about Bhartṛhari's notion of 'metaphorical existence' assignable to what we call fictions. The last point helps to explain the negative existential statements and other related problems. Chakrabarti concludes with a brief note on Gaṅgeśa, the Navya-Nyāya author, who insisted that although the 'unnegatable' properties (i.e. the universal properties as defined above) are acceptable in the Nyāya system and the 'unlocatable' properties (fictions) are unacceptable, for we can not say that absence or lack of such fictions characterizes each reality and thereby only turn the lack of a fictional property into an acceptable 'universal' property. Gaṅgeśa argued that to talk about the absence of fictional properties would amount to admission of fictional properties into the system and hence a better way would be to analyse fictions, after Udayana, into their components and then talk about the lack of connection between such components. The assumption is that fictional properties are always composite properties. This may well coincide with the intuition of some modern analytical philosophers.

2.14 Reference, Sense and Nyāya

Shaw gives first an account of Frege' doctrine of SENSE and then goes on to note the salient features of Kripke's New Theory of Reference in order to prepare ground for his exposition of what he calls the Nyāya theory of proper names. He in fact refers to several views of proper names propounded by different Naiyāyikas at some time or other. He tries to underline, whenever possible, the apparent similarity between different aspects of the Nyāya view and the views ascribable to Frege, Russell and Kripke.

In order to expound the standard Nyāya doctrine about how the utterance of words generates our awareness (the hearer's awareness) of the objects taken to be their 'meanings' (artha), Shaw briefly refers to the classification of words (those appearing in actual sentences) into two sets: those

having primary (lexical) meaning and those having secondary (metaphorical) meaning. He also refers to the notion of ākāṅkṣā or syntactic expectancy, the service of which is required in Nyāya to account for the relatedness of objects or meanings (see Siderits for other alternative approaches to the same problem) in the object-complex or the structural content of the cognition generated by the utterance of a sentence-token. If by "meaning" of a sentence we mean this 'whole', this particular structural content of such a cognitive episode, then it becomes easy for us to say with Nyāya that although the component words individually contribute to the formulation of the whole sentence-meaning, the actual sentence-meaning is a separate whole and not simply the meanings of the words added up, for the connectedness of such elements is supplied by ākāṅkṣā. Siderits has given the Prābhākara-Bhāṭṭa controversy on this matter.

To return to proper names. Shaw derives a lot of leverage out of the 'epistemic overtones' of Frege's sense-theory as well as its 'mode of presentation' element. This apparently enables Shaw to compare Nyāya with Frege, and in this he is not very far from what was suggested (very impressionistically) by Bhattacharya (S) in his "mode of presentation" analysis of the Fregean sense (in an earlier publication). Shaw argues that the Nyāya view of proper names can be interpreted in the context of Nyāya ontology and epistemology (cf. ultimate differentia, uniqueness) in such a way as would underline many similarities between Nyāya and certain modern views about proper names. Particularly he mentions that the 'name-giving ceremony' as well as the 'intentional' aspect of Kripke's view of proper names are points of such comparison. Besides, he insists that the epistemic aspect of the Nyāya theory (the notion of delimitor etc.) has its parallel in the sense-theory of Frege.

2.15 Awareness and Meaning

Matilal tries to explain the Nyāya epistemic principle of awareness by which Navya-Nyāya resolves particularly the problem of an "epistemic first" as a pre-condition for a judgemental or qualificative awareness (an awareness of an a as an F). It is claimed that an awareness of an a as an F, or an a as F (or an awareness that a is F) presupposes a prior conception of what is meant by "F". Such a conception may be a simple awareness of what it is to be F, or to be more articulate, an awareness of F-ness, or even an awareness of f if the awareness is verbalizable as "a has f"). Our theory assumes that the verbal form of our basic verbalizable awareness is "a is F" although it is not necessry for it to be actually verbalized. The general structure of this basic awareness is: a has f or a (is) f-possessing. We can

generally describe this basic awareness as that of an _a_ as _F_. A perceptual knowledge of something as _F_ ('white' for instance) would therefore presuppose under this theory a conception (awareness) of F-ness.

Since an awareness or a piece of knowledge is treated as an episode arising in a subject (in this theory), the above requirement is causal. But the connection may be called also 'logical' in the sense that one awareness-type (of which the conception or awareness of F-ness is a token) causes the other awareness-type (the qualificative awareness). When I perceive something as an elephant there must have been a prior awareness of elephanthood in me (consult K. Bhattacharya's example below). This prior awareness may be given by memory but in most cases of perception (where it is difficult or even impossible to imagine the presence of such a memory to be revived, or where supposition of memory-revival would be too complicated to be real) it is simpler to suppose that a _simple_ perception (awareness) of elephanthood has "unconsciously" taken place in the subject prior to his perception of that thing as an elephant. (Even our dispositionality to recognize _x_ as an elephant has to be acquired, according to empirical philosophy, at a given perceptual situation.)

The above principle of awareness is generalized in Nyāya to tackle many related problems such as that of our (the hearers') awareness of an object arising from the utterance of a proper name "Socrates" or any other name such as "water". It is assumed (by most) that unlike perception our verbal awareness, i.e. awareness arising from the utterance, cannot be a _simple_ awareness of an object _per se_. Hence when a proper name "Socrates" is uttered the hearer becomes aware of a person/object which should be presented in his awareness under a 'mode' of presentation. In other words the hearer's awareness must have the structure: _a_ as _F_. But under what _mode_ is the person/object Socrates presented in our awareness when the name is uttered and heard by us? Different Naiyāyikas suggest slightly different answers. But one common factor among them is that a 'distinguisher' must act as the said _mode_ (see also Shaw below). It should also be noted here that Nyāya does not subscribe to representationalism and hence the presentation of the object in the awareness should not be taken to imply a third realm of 'objects' between the external world outside and the awareness inside.

If by the 'referent' of a term we mean the object (as distinguished from the mode or distinguisher) that is presented to the hearer's awareness generated by the relevant utterance, then such a referent 'floats' in the said awareness always being distinguished by a distinguisher. (If we say that this distinguisher picks out the object for the hearer then this has a distant resonance of the sense-theory of Frege). Such a distinguisher may be essential and relevant or

it may be arbitrary. An essential distinguisher may be part of the 'meaning-complex' of the word but an arbitrary distinguisher is like an 'unmeaning mark' whose purpose is served as soon as the object is singled out for the hearer. In other words, the hearer may pick out the object through any arbitrary property derived from the body of information available to him in the context as long as it uniquely distinguishes the object in question, but such distinguishers need not be part of the 'meaning-complex' of the word. Gadādhara says that in certain indexicals such as "it" the arbitrary distinguisher may be the property of being in the thought of the speaker, but once the relevant object (say, a chair) is picked out in this way, it would 'float' in the awareness under the mode of chairhood, the so-called essential and most familiar attribute of the object.

3.

In the above I have indicated how often some of the issues discussed in this volume can be further explored in what we have called 'comparative perspective'. The notions of knowledge, belief, intension, intention and action are all relevant to a large extent to modern philosophical discussion, and it is expected that our comparative perspective will only enrich such discussion. In particular I believe that both our philosophy of language and philosophy of mind will be only more substantial if we extend our horizon to include the discourses of Nyāya, Mīmāṃsā and Buddhism.

The problems of proper names, universals, intentionality and propositional attitudes are live issues in certain circles of present day philosophy. Insights of the classical Indian philosophers have been shown to be helpful in throwing light upon some of these questions. The New Theory of Reference, for example, has opened up many sides of the familiar problem of meaning. In many ways it supercedes the traditional approach, but the insights of traditional philosophers as well as classical Indian philosophers need not always be dispensable. On the problem of universals in particular, the Buddhist developed the theory of apoha or 'exclusion of others' which is not represented in this volume (see forthcoming B.K. Matilal (ed.) Buddhist Logic and Epistemology, D. Reidel). A proper assessment of this theory in the context of modern developments has not yet been done and hence it could very well be part of the future undertaking of the research workers in this interdisciplinary area (consisting of the analytical aspect of Indian philosophy and modern philosophical analysis as it is understood today in the West). Bhartṛhari's sentence holism is another very intriguing doctrine, which once attracted the attention of W.V. Quine (see Word and Object, p. 9). Both A. Chakrabarti and K. Bhattacharya have briefly

dealt with Bhartṛhari's holism. A proper study and analysis of Bhartṛhari's linguistic insights can provide us with an area of fruitful research. Among other things, his sentence-holism can be linked up with the discussion of the syncategorematic treatment of predicates that has been advocated by Quine (see P. Gochet below). There are many other areas where such interdisciplinary studies can be undertaken. As far as the classical Indian material is concerned, we have only "scratched the surface". Properly annotated translations of main texts and studies are long overdue. We hope the papers presented here will help to open up a fertile field of philosophical research.

The chief purpose here in this volume has been an attempt to initate a dialogue between the ancient (Sanskrit classical) philosophers and the modern philosophers--a dialogue as much as it is possible and can be allowed in the pages of an anthology. Too often Indian philosophy has been considered (very wrongly) as being very 'soft' and tender-minded. Too often it has been identified as being mystical, non-argumentative, poetic and dogmatic. An emphasis on the other side has been attempted here to correct this heavily one-sided picture. What best way is there to accomplish this than by initiating eventually a dialogue with modern analytical philosophers in a way that would try to transcend the language barrier? This volume is a positive step in this direction.*

---

*. I wish to thank Professor L. Linsky, Professor P.K. Sen, and Professor J.N. Mohanty who read some portions of the earlier version of this introductory essay.

M.J. Cresswell

WE ARE ALL CHILDREN OF GOD

In a recent article, 'An Immaculate Conception of Modality', Brian Skyrms [14] has shewn how to interpret the necessity operator in various modal systems as a metalinguistic predicate of sentences. The principal task of the present paper will be to shew that any intensional operator may be construed in this way. I will then make some remarks about the philosophical importance of this fact, the burden of which will be that it frees a philosopher to do things intensionally or metalinguistically according to taste. I will then link the equivalence proofs with some suggestions Robert Stalnaker has made in connection with the problem of the semantics of propositional attitudes. The paper is self-contained but I suspect that a better appreciation of what is going on will be aided by a careful reading of Skyrms' paper. Also, some familiarity with Stalnaker's use of double indexing in, e.g., [16] will help in the later sections. I would like to thank Steve Boër who drew my attention to the Skyrms article when he was in Wellington in July and August of 1979.

1. SKYRMS' RESULT

In this section I will set out Skyrms' result for S5 and prove it in a form which will allow easy generalization to my principal theorems.[1]

We begin with a language L_0 [14, p. 369] which is interpreted by a class $\mathfrak{M}_0$ of models (Skyrms does not give this class a name; it will emerge in a moment that it is important to do so). Skyrms is very liberal in deciding what can count as a model. A model is in fact any structure which assigns truth values to formulae of L_0. I shall use $M \models \alpha$ to mean that α is true in the model M and $M \dashv \alpha$ to mean that α is not true in the model M. The models are all assumed to be bi-valent but otherwise subject only to the condition that they respect the truth functors. In section two I shall drop this latter requirement, but in section one, in the exposition of Skyrms, I shall retain it.

Given then our class of models $\mathfrak{M}_0$, indexed by some set I, we have it that each model M_0^i in $\mathfrak{M}_0$ gives to every sentence of L_0 a truth value. Based on L_0 we erect languages of two kinds, first a modal language L_M; and second a family of languages L_n which combine to produce a language L_ω.

B. K. Matilal and J. L. Shaw (eds.), Analytical Philosophy in Comparative Perspective, 39–60.

L_M is simply the language of modal logic whose atomic sentences are the sentence of L_0.

L_n ($n \geq 0$) is a sequence of languages whose formation rules are as follows:

(i) If α is a sentence of L_n then α and $L\alpha$ are sentences of L_{n+1}.

(ii) If α and β are sentences of L_n then so are $(\alpha \vee \beta)$, $(\alpha \wedge \beta)$, $(\alpha \supset \beta)$, $\sim\alpha$ and $\sim\beta$.

In these conditions I have departed from Skyrms' notation in two ways. First I have used α and β in place of his S and T, and second I have used only a single symbol L in place of Skyrms' * and Q. Q(S) for Skyrms represents in L_{n+1} the name of the sentence S from L_n. But his formation rules require that * and Q always occur together and his model theory on p. 370 requires that they be evaluated together. *, for Skyrms, means 'is valid'. (Any reader who wishes may take L as an abbreviation for the two symbols taken together.)

C is a mapping from the sentences of L_M to the sentences of L_ω such that:

(i) If α is free of modalities then $C(\alpha)$ is α.

(ii) If α is $\Box\beta$ then $C(\alpha)$ is $L(C(\beta))$.

(iii) If α is $(\beta \vee \gamma)$, $(\beta \wedge \gamma)$, $(\beta \supset \gamma)$, or $\sim\beta$ then $C(\alpha)$ is, respectively, $(C(\beta) \vee (C(\gamma))$, $(C(\beta) \wedge (C(\gamma))$, $(C(\beta) \supset (C(\gamma))$, $\sim C(\beta)$.

What it comes down to is that a formula of L_M has $\Box$ in those and only those places where a formula of L_ω has L. It would

in fact be possible to do the whole thing in one language and let the difference reside entirely in the interpretations. L_M will be interpreted in the usual way using a set of possible worlds. L_ω on the other hand will be given an interpretation

of a metalinguistic kind which may be defined as follows:

We are given a family of models, $\mathfrak{M}_0$ for L_0 indexed by I. We shew how each $M_0{}^i$ ($i \in I$) in $\mathfrak{M}_0$ induces a sequence of models $M_n{}^i$ for every L_n in such a way that every wff α of L_n is either true or false in $M_n{}^i$. (Skyrms speaks as if a single model for L_0 induced such a sequence on its own. As will become apparent, the models induced at higher levels depend on the whole of $\mathfrak{M}_0$. We can let $\mathfrak{M}_n$ denote all the $M_n{}^i$ ($i \in I$).)

$M_n{}^i{}_{+1}$ is defined as follows:

(i) If $\alpha \in L_n$ then
$M_n{}^i{}_{+1} \models \alpha$ iff $M_n{}^i \models \alpha$

(ii) If α is $L\beta$ and $\beta \in L_n$ then
$M_n{}^i{}_{+1} \models \alpha$ iff $M_n{}^j \models \alpha$
for every $j \in I$.

(Note how (ii) makes reference to other n-level models besides

M_n^i, and so M_n^{i+1} could be changed even though M_n^i is not.)

M_ω^i is the union of all the M_n^i's. It is a model for L_ω.

$\mathfrak{M}_\omega$ is the collection of all the M_ω^i's.

It should not be too hard to see that Skyrms' A and B on p. 370 correspond exactly to this definition when we read 'valid' as 'true in all models'.

Skyrms is now able to prove that there is an intimate connection between the formulae valid by this semantics and the theorems of propositional S5. In order to establish his result in a way that can be used in later sections, I shall prove two theorems.

I will assume that an S5 model for L_M is an ordered pair $\langle W,V \rangle$ in which W is a set (of "possible worlds") and V is an assignment which assigns to each atomic sentence a subset of W.[2] We define what it is for a formula of L_M to be true in a world $x \in W$ as follows:

(i) If α is an atomic sentence then $\langle W,V \rangle \models_x \alpha$ iff $x \in V(\alpha)$

(iia) If α is $\beta \vee \gamma$ then $\langle W,V \rangle \models_x \alpha \vee \beta$ iff $\langle W,V \rangle \models_x \alpha$ or $\langle W,V \rangle \models_x \beta$ (and so on for $\wedge$ and $\supset$).

(iib) If α is $\sim\beta$ then $\langle W,V \rangle \models_x \alpha$ iff $\langle W,V \rangle \dashv_x \beta$ ($\langle W,B \rangle \dashv_x \beta$ just means that it is not the case that $\langle W,V \rangle \dashv_x \beta$)

(iii) If α is $\Box\beta$ then $\langle W,V \rangle \models_x \Box\beta$ iff $\langle W,V \rangle \models_y \beta$ for every $y \in W$.

In order to state these theorems I will impose an extra condition on L_0 (much as Skyrms does on p. 373). I shall insist that no sentence of L_0 is a truth functional complex. The point of this is simply to make it possible to have a class $\mathfrak{M}_0$ of the models for L_0 which assigns to them any truth values whatever. If they were truth functional complexes the fact that $\mathfrak{M}_0$ respects truth functors could prevent this.

THEOREM 1.1 If $\langle W,V \rangle$ is an S5 model for L_M then there is a class of models $\mathfrak{M}_0$ for L_0 indexed by W such that for any wff α of L_M where $C(\alpha)$ is a wff of L_n then

$$\langle W,V \rangle \models_x \alpha \text{ iff } M_n^x \models C(\alpha)$$

$\mathfrak{M}_0$ is defined as follows:
where α is an atomic sentence of L_M then α will be a sentence of L_0 and $C(\alpha) = \alpha$. Let $M_0^x \models \alpha$ iff $x \in V(\alpha)$. Then it is clear that the theorem holds for n = 0. But for a wff of L_0

then $M_n{}^x \models \alpha$ iff $M_0{}^x \models \alpha$ and so the theorem holds for atomic formulae.

Suppose that α is $\sim\beta$ and that the theorem holds for β. Now iff $\beta \in L_n$ then $\alpha \in L_n$. So

$$\begin{aligned} \langle W,V\rangle \models_x \alpha \text{ iff } & \langle W,V\rangle \dashv_x \beta \\ \text{iff } & M_n{}^x \dashv C(\beta) \\ \text{iff } & M_n{}^x \models \sim C(\beta) \\ \text{iff } & M_n{}^x \models C(\sim\beta) \end{aligned}$$

The induction for the other truth functors is exactly analogous. Suppose α is $\Box\,\beta$. Then $\alpha \in L_n$ where $n > 0$ and $\beta \in L_{n-1}$.
So

$$\langle W,V\rangle \models_x \alpha \text{ iff } \langle W,V\rangle \models_y \beta$$

for every $y \in W$; i.e. iff $M_n{}^y{}_{-1} \models C(\beta)$
for every $y \in W$; i.e. iff $M_n{}^x \models L(C(\beta))$
iff $M_n{}^x \models C(\Box\,\beta)$
iff $M_n{}^x \models C(\alpha)$

THEOREM 1.2 If $\mathfrak{M}_0$ is a class of models for L_0 indexed by a set I and L_M is a modal language based on L_0 then there is an S5 model $\langle I,V\rangle$ such that for any wff α of L_M and any $i \in I$,

$$\langle I,V\rangle \models_i \alpha \text{ iff } M_n{}^i \models C(\alpha)$$

where $C(\alpha) \in L_n$.

The proof is just like that of Theorem 1. The "possible worlds" of the S5 model are just the members of the index set of $\mathfrak{M}_0$ and V is defined in the obvious way: viz. where $\alpha \in L_0$ (in this case $C(\alpha) = \alpha$), $i \in V(\alpha)$ iff

$$M_0{}^i \models \alpha$$

The inductive proof then proceeds as before.

We now have Skyrms' soundness proof on p. 371 as an immediate consequence of Theorem 1.2, using the soundness of S5 with respect to the standard modelling. For if $C(\alpha)$ is not valid in L_ω then there is a family $\mathfrak{M}_\omega$ of models in one of which it fails, so by Theorem 1.2 there is an S5 model in which α fails and so α is not a thesis of propositional S5. This is Theorem 1.3:

THEOREM 1.3 [14, p. 371] For any choice of L_0, if α is a thesis of propositional S5 then $C(\alpha)$ is true in all models of L_ω.

We can also get a converse which is more or less Skyrms' completeness theorem on p. 374. Consider L_M to be the language of propositional modal logic obtained by treating the

sentences of L_0 as if they were propositional variables. (Remember they have no truth-functional complexity.)

THEOREM 1.4 For any wff a of L_M, if $C(a)$ is valid in all models of L_ω then a is a theorem of propositional S5.

Suppose that a is not a theorem of propositional S5. Then a fails in an S5 model $\langle W,V\rangle$ in which the sentences of L_0 are treated as atomic sentences. So by Theorem 1.1 $C(a)$ fails in a model of L_ω.

2. PROPOSITIONAL LANGUAGES

I want to shew in this section that Skyrms' results can be extended to any propositional language. By a propositional language I mean a language of the kind considered in chapter one of [5].[3] It is a language based on propositional letters and an assortment of n-place propositional connectives for various n. If $\mathfrak{L}$ is a propositional language then the nature of the propositional letters is completely irrelevant. From the point of view of $\mathfrak{L}$ and its models they are simple, but nothing prevents them from having a complex internal structure of their own, and in particular we can regard them as sentences of a language L_0 (which may not itself be a propositional language) and so preserve the connection with Skyrms' work.

A propositional language $\mathfrak{L}$ then will consist of a set L_0 of atomic sentences (atomic that is with respect to $\mathfrak{L}$) and, for each n, a set of Δ_n of n-place propositional functors. The formation rules are as follows:

> (i) If $a \in L_0$ then a is a sentence of $\mathfrak{L}$ (ii) If $\delta \in \Delta_n$ (i.e. if δ is an n-place propositional functor) and $a_1, \ldots, a_n$ are sentences of $\mathfrak{L}$ then $\delta a_1 \ldots a_n$ is also a sentence of $\mathfrak{L}$.

A propositional language $\mathfrak{L}$ can be given an intensional interpretation in the following way: we define what I called in [5, p. 25] an <u>indexical</u> <u>model</u> (but will here call an <u>intensional</u> model) for $\mathfrak{L}$ as a pair $\langle W,V\rangle$ in which W is a set of indices (this seems the most neutral term but they can be called possible worlds). Unlike the S5 models of the previous section, V assigns values to all the symbols of $\mathfrak{L}$ and not just the atomic sentences. V is in fact a function such that

> (i) If $a \in L_0$ then $V(a) \subseteq W$
> (ii) If $\delta \in \Delta_n$ then $V(\delta)$ is a function from n-tuples of subsets of W to subsets of W.

(i) means that, as in section one, atomic sentences are

true or false in various worlds and (ii) means that an n-place propositional functor tells us which worlds the complex it makes is true in on the basis of which worlds the sentences which are its arguments are true in. This can be made precise by giving a truth definition for members of Ω .

(i) If $\alpha \in L_0$ then $\langle W,V \rangle \models_x \alpha$ iff $x \in V(\alpha)$
(ii) If α is $\delta\alpha_1 \ldots \alpha_n$ where δ is an n-place functor then

$$\langle W,V \rangle \models_x \delta\alpha_1 \ldots \alpha_n \text{ iff } x \in V(\delta)\,(|\alpha_1|, \ldots , |\alpha_n|)$$

where

$$|\alpha_i| = \{x \in W : \langle W,V \rangle \models_x \alpha_i\}$$

(i.e. $|\alpha_i|$ is just the set of worlds at which α_i is true).

The next task is to express all this metalinguistically as Skyrms did for S5. First we need to set out the sequence L_ω of languages. In S5 the truth functors did not raise the level of the language. In the general case <u>all</u> functors raise the level of the language whether they turn out to be truth functors or not.

In this section I shall follow a procedure mentioned in section one, and shew how the very same language Ω itself can be set out as a hierarchy of languages and be given a metalinguistic interpretation. We could if we wished express each functor δ in the hierarchical language as a combination of the functor and a quotation operator, but as in the modal case there is nothing to be gained by so doing.

The set of sentences of L_{n+1} is in fact the smallest set S such that if $\alpha_1, \ldots , \alpha_k$ are in sentences of L_n and $\alpha \in \Delta_k$, then $\alpha_1, \ldots , \alpha_k$ and $\delta\alpha_1 \ldots \alpha_k$ are in S. Where L_ω is the union of all the L_n's, then $L_\omega = \Omega$.

We now imagine that we have a set $\mathfrak{M}_0$ of models for L_0. We want to use $\mathfrak{M}_0$ to induce a family of models for L_ω. In order to do this we must interpret the propositional functors. Each functor δ is to be associated with a certain property of models. Strictly speaking, where δ is a k-place functor, it will be associated with a k + 1 relation R between a model and a class of models. It must however be a relation which is stratified in a certain way so that it only relates models of the same language level. In other words we require that the k+1-place relation R associated with a k-place functor δ be a relation between a model M and classes of models $A_1, \ldots , A_k$ such that where $MR\langle A_1, \ldots , A_k\rangle$ and M is a model for L_{n+1} then $A_1, \ldots , A_k$ are all classes of models for L_n.

Given our class $\mathfrak{M}_0$ of models for L_0 indexed by I we can now define the class $\mathfrak{M}_n$ of models for L_n induced by $\mathfrak{M}_0$ in the following way:

In this definition where a is a formula for L_n we use the notation $|a|_n$ for

$$\{M : (\exists i \in I)\ (M = M_n^i \text{ and } M_n^i \models a)\}$$

(i.e. $|a|_n$ is the class of L_n models, from the set generated by $\mathfrak{M}_0$, which make a true). What we shew is how to define an M_{n+1}^i based on M_n^i as in the previous section.

In fact the definition is simpler than it was in section one because we only have to shew how to define M_{n+1}^i for a formula a when a is $\delta a_1 \ldots a_k$ and $a_1, \ldots, a_k$ are all in L_n. The definition is:

(i) If $a \in L_n$ then $M_{n+1}^i \models a$ iff $M_n^i \models a$

(ii) $M_{n+1}^i \models \delta a_1 \ldots a_k$

iff

$M_{n+1}^i\ R\langle |a_1|_n, \ldots, |a_k|_n\rangle$

Each M_ω^i is the union of all the M_n^i's, and $\mathfrak{M}_\omega$ is the class of all the M_ω^i's.

Before proceding to shew that this semantics is equivalent to the intensional one, it might be worth observing one or two things. In the first place R can be taken extensionally so that the properties of models that we are interested in are purely extensional properties. In the second place, the property of validity used in Skyrms' proof can be easily seen as a special case of R. R is in fact the validity relation (relative to M_ω) when it is a two-place relation between a model in $\mathfrak{M}_{n+1}$ and a class A of models in $\mathfrak{M}_n$ such that M_{n+i}^i R A iff A = $\mathfrak{M}_n$. (I.e. it will be true, in any model in $\mathfrak{M}_{n+1}$ that a formula of L_n is valid, relative to $\mathfrak{M}_n$, iff the class of L_n models in which it is true is the whole of $\mathfrak{M}_n$; i.e. iff it is true in every model in $\mathfrak{M}_n$.)

A set of models $\mathfrak{M}_0$ for L_0 thus induces a set of models $\mathfrak{M}_\omega$ for L_ω(= $\mathfrak{L}$) with the aid of a set $\mathfrak{R}$ of relations associated with the propositional functors of $\mathfrak{L}$. We shew how this kind of semantics is equivalent to the intensional one.

THEOREM 2.1 If $\langle W,V\rangle$ is an intensional model for a propositional language $\mathfrak{L}$ based on a set L_0 of atomic sentences then there is a class $\mathfrak{M}_0$ of models for L_0 indexed by W, and a

class $\mathfrak{R}$ of properties of models such that each $R \in \mathfrak{R}$ is associated with a functor δ, and the models $\mathfrak{M}_\omega$ are such that for any wff a of $\mathfrak{L}$

$$\langle W,V\rangle \models_x a \text{ iff } M_n{}^x \models a$$

where $a \in L_n$.

We construct $\mathfrak{M}_\omega$ in the following way. Where a is an atomic sentence we simply define $M_0{}^x \models a$ iff $x \in V(a)$. Since an atomic sentence $a \in L_n$ for every n then the theorem holds for atomic sentences of $\mathfrak{L}$. We suppose that it holds for $a_1, \ldots, a_k$ in L_n and consider a k-place functor δ and the wff $\delta a_1 \ldots a_k$ in L_{n+1}. To deal with this case we must define the relation R. Now R may be a relation between modals of any level but it suffices to assume that it has already been defined for classes of models for $L_{m<n}$ and that it will later be defined for classes of models for $L_{m>n}$.

Let R be the following relation:
R holds between $M_n{}^x{}_{+1}$ and the classes $A_1, \ldots, A_k$ of models of L_n iff there are wff $a_1, \ldots, a_k$ of L_n such that $A_1, \ldots, A_k$ are respectively $|a_1|_n, \ldots, |a_k|_n$ and

$$\langle W,V\rangle \models_x \delta a_1 \ldots a_k{}^4$$

We must shew that

$$\langle W,V\rangle \models_x \delta a_1 \ldots a_k \text{ iff } M_n{}^x{}_{+1} \models \delta a_1 \ldots a_k$$

Now $M_n{}^x{}_{+1} \models \delta a_1 \ldots a_k$ iff

$$M_n{}^x{}_{+1}\ R\ \langle |a_1|_n, \ldots, |a_k|_n\rangle$$

where R is the relation associated with δ. But by definition of R

$$M_n{}^x{}_{+1}\ R\ \langle |a_1|_n, \ldots, |a_k|_n\rangle \text{ iff}$$

there are some $\beta_1, \ldots, \beta_k$ in L_n such that $|\beta_1|_n = |a_1|_n, \ldots, |\beta_k|_n = |a_k|_n$, and $\langle W,V\rangle \models_x \delta\beta_1 \ldots \beta_k$.
By the induction hypothesis where $1 \leq h \leq k$ for any $y \in W$:

$$M_n{}^y \models a_h \text{ iff } \langle W,V\rangle \models_y a_h$$

and

$$M_n{}^y \models \beta_h \text{ iff } \langle W,V\rangle \models_y a_h$$

so

$$\langle W,V\rangle \models_y a_h \text{ iff } \langle W,V\rangle \models_y \beta_h$$

so

$$\langle W,V\rangle \models_x \delta a_1 \ldots a_k \text{ iff } \langle W,V\rangle \models_x \delta\beta_1 \ldots \beta_k$$

so this means that $M_n{}^x{}_{+1}$ R $\langle |a_1|, \ldots, |a_k| \rangle$
iff $\langle W,V \rangle \models_x \delta a_1 \ldots a_k$.

In other words

$$M_n{}^x{}_{+1} \models \delta a_1 \ldots a_k \text{ iff } \langle W,V \rangle \models_x \delta a_1 \ldots a_k$$

as required by the induction.

THEOREM 2.2 Suppose that $\mathfrak{M}_\omega$ is a class of models for L_ω based on a class of models $\mathfrak{M}_0$ for L_0 (indexed by I). Then there is an intensional model $\langle I,V \rangle$ such that for any wff a of L (= $\mathfrak{L}$) and any i $\in$ I

$$\langle I,V \rangle \models_i a \text{ iff } M_n{}^i \models a$$

where $a \in L_n$.

All we need to do is to define V. If $a \in L_0$ then $i \in V(a)$ iff $M_0{}^i \models a$. If δ is a k-place functor then $V(\delta)$ is defined as follows:

> Let $i \in I$; and let $a_1, \ldots, a_k$ be subsets of I; then $i \in V(\delta)(a_1, \ldots, a_k)$ iff there exist wff $a_1, \ldots, a_k$ in some L_n such that $a_h = \{i \in I : M_n{}^i = a_h\}$ $(1 \le h \le k)$ and
>
> $$M_n{}^i{}_{+1} \models \delta a_1 \ldots a_k$$

Obviously the inductive proof of the theorem holds for the atomic sentences. We must shew that if it holds for $a_1, \ldots, a_k$ then it holds for $\delta a_1 \ldots a_k$. (In the proof that follows I shall use the notation $|a|$ as an abbreviation for $\{i \in I : \langle I,V \rangle \models_i a\}$ and $|a|_n$ as an abbreviation for $\{M \in \mathfrak{M}_i : (\exists i)(M = M_n{}^i \text{ and } M_n{}^i \models a)\}$.

Now $\langle I,V \rangle \models_i \delta a_1 \ldots a_k$

iff $i \in V(\delta) \langle |a_1|, \ldots, |a_k| \rangle$

i.e. iff there is some $\beta_1, \ldots, \beta_k$ in some L_m such that

$$|a_h|_n = |\beta_h|_m \ (1 \le h \le k)$$

and $M_m{}^i{}_{+1} \models \delta\beta_1 \ldots \beta_k$
(where $\delta\beta_1 \ldots \beta_k \in M_m$).

In order to prove the inductive step then we have to establish that where $|a_h|_n = |\beta_h|_m$
then $M_n{}^i{}_{+1} \models \delta a_1 \ldots a_k$ iff $M_m{}^i{}_{+1} \models \delta\beta_1 \ldots \beta_k$
let j = max (n,m)
then (by the construction of $\mathfrak{M}_\omega$)

$$|a_h|_j = |\beta_h|_j$$

So $M_j{}^i{}_{+1}$ R $\langle |\alpha_1|_j, \ldots, |\alpha_k|_j\rangle$ iff $M_j{}^i{}_{+1}$ R $\langle |\beta_1|_j, \ldots, |\beta_k|_j\rangle$,

So $M_j{}^i{}_{+1} \models \delta\alpha_1 \ldots \alpha_k$ iff $M_j{}^i{}_{+1} \models \delta\beta_1 \ldots \beta_k$
But, by construction of $\mathfrak{M}_{j+1}$

$$M_j{}^i{}_{+1} \models \delta\alpha_1 \ldots \alpha_k \text{ iff } M_n{}^i{}_{+1} \models \delta\alpha_1 \ldots \alpha_k$$

and

$$M_j{}^i{}_{+1} \models \delta\beta_1 \ldots \beta_k \text{ iff } M_m{}^i{}_{+1} \models \delta\beta_1 \ldots \beta_k.$$

So

$$M_n{}^i{}_{+1} \models \delta\alpha_1 \ldots \alpha_k \text{ iff } M_m{}^i{}_{+1} \models \delta\beta_1 \ldots \beta_k$$

as required.

This establishes the inductive step, and therefore proves the theorem. Theorems 2.1 and 2.2 are the principal results of the paper.

Let us suppose now that there is a collection of sentences of L_ω which form a logic Λ in the sense that there is a class of models $\mathfrak{M}_\omega$ in which precisely the members of Λ are valid. What theorem 2.2 shews is that there is an intensional model in which precisely the members of Λ are true in every world. Conversely Theorem 2.1 shews that if there is an intensional model (like the canonical models of modal logic) in which precisely the members of Λ are true in all worlds then there is a class $\mathfrak{M}_\omega$ of metalinguistic models which precisely define validity in Λ. (Conditions under which a set Λ of sentences will be a logic in this sense will be discussed in section four.)

3. LOGICAL CONSTANTS

There is what appears at first sight to be a very important difference between what was done in section one and what was done in section two. In section one it was shewn how a collection of models $\mathfrak{M}_0$ for L_0 induced a unique collection $\mathfrak{M}_\omega$ of models for L_ω. In section two $\mathfrak{M}_\omega$ depended not merely on $\mathfrak{M}_0$ but in addition on the relations associated with each propositional operator.

The difference has already been shewn to be somewhat illusory because we saw that Skyrms' result could be expressed as involving a special case of a relation R. In section one the definition of how $\mathfrak{M}_0$ induced $\mathfrak{M}_\omega$ incorporated the particular meaning that was to be given to L, for all that it did not make specific reference to any R. It might of course be

objected that the property of validity is a highly general property which is defined for every collection of models, whereas the relations R may well depend on all kinds of facts about the models, possibly very specific facts; and might not even make sense for all but a few models. This may all be true but it does not affect the question of whether an operator δ is treated intensionally or metalinguistically. What it does allude to is perhaps the special nature of the operator L(□) however treated, whether intensionally or metalinguistically; and maybe marks off features which distinguish it from most other operators. Perhaps it is that L is a logical constant. I am personally rather sceptical about whether there really is a distinction between logical words and non-logical words, so I don't feel obliged to say what it is to claim that L is a logical constant. However, if the distinction can be drawn, then I suggest it would be because of that (that it is a logical constant) that section one has a different flavour from section two, where a metalinguistic interpretation was given for every functor whether a logical word or not.

One respect which might make a difference is the fact that R has to be a relation. Whether or not $M_n{}^i \models \delta a_1 \ldots a_k$ depends on i, as well as on the truth of $a_1, \ldots, a_k$ in the models of $L_{m<n}$. In section one the truth of La depended on the truth of a in the lower level models but was independent of the choice of i. This meant that L had to be construed as an S5 type of necessity in which any fully modalized sentence is true either in all worlds or in none. In weaker modal systems (about which I shall say something in just a moment) a formula □a might be true in one world but not true in every world. In intensional semantics these systems are studied by imposing an accessibility relation on the set of worlds. One might perhaps say that the need for this extra relation prevents the □ in such systems from being a genuinely logical word.[5] If so then what distinguishes section one from section two is simply the fact that in section one we were interested only in a property of a class of models (or a k-tuple of classes) of models instead of a relation between a model and a class (k-tuple of classes) of models.

Actually the situation is even more complicated because in section one the truth functors were not treated metalinguistically. If they were they could not be treated absolutely but would need a relation R which behaved selectively with repect to different models. No doubt the question is an interesting one in its own right but it is perhaps a convenient time to recall that this issue is independent of whether the operators are treated intensionally or metalinguistically. Intensionally these absolute models are simply ones where an operator forms an expression which if true in any world is true in every world.

I mentioned a little while ago that the modal systems weaker than S5 are treated (at least in the case of those systems which are called 'normal' modal systems [12, p. 12] by adding to the (intensional) models a relation of 'accessibility' between worlds. The appropriate truth clause then becomes (where R is the accessibility relation)

$$\langle W,R,V\rangle \models_x \Box a \text{ iff } \langle W,R,V\rangle \models_y a$$

for every y such that xRy.
By imposing various conditions on R, we obtain the different modal systems. Put metalinguistically, we would impose an accessibility relation on the models of $\mathfrak{M}_0$ and then would define

$$M_n{}^{i}{}_{+1} \models La \text{ iff } M_n{}^{j} \models a$$

for every j such that $M_0{}^{i}$ R $M_0{}^{j}$.

An accessibility relation R can always be expressed as a 'neighbourhood' relation R* by defining R* so that for any $x \in W$, and $A \subseteq W$, x R* A iff $\{y \in W : xRy\} \subseteq A$ ([4, p. 352] and [12, p. 22]). Obviously a metalinguistic accessibility relation corresponds in the same way to a metalinguistic neighbhourhood relation (more precisely: $M_n{}^{i}{}_{+1}$ R* A iff where $M_0{}^{i}$ R $M_0{}^{j}$ then $M_n{}^{j} \in A$) and so all the results of section two carry over to all the normal modal logics.

Of course this result does not tell us how to interpret R so that it makes any reasonable intuitive sense. Different conditions lead to different modal systems. Skyrms spends the second part of his paper shewing how to get S4 out of a 'provability' interpretation for the necessity operator. Put in terms of the present paper what Skyrms is doing may be seen as shewing how to give R some intuitive content when it is a relation between models connected with the notion of provability. The moral here would seem to be that whether one chooses an intensional or a metalinguistic interpretation is a question of elegance and convenience. It cannot be decided on formal grounds.

The results of this section can probably be extended to more complex languages like the categorial languages discussed in chapter five of [5] and elsewhere, but their precise statement would need great care because of the possibility of functor expressions which are complex and have to have a meaning which is determined by the meanings of their parts and cannot be simply a property "associated" with them. Although languages of this kind have a very straightforward intensional semantics, it is doubtful, to me at least, whether any intuitively reasonable metalinguistic interpretations would be forthcoming. The best approach here for one who is enamoured of the metalinguistic interpretations would seem to be to

translate these languages into a more tractable intensional language (one, say, of first-order) and then impose the metalinguistic interpretation.

4. MAKING PIGS FLY

We know that there are two (at least) ways of making the sentence

(1) 'pigs fly'

true. One is to make pigs fly and the other is to begin to use 'pigs' the way we now use 'birds'. This difference can be easily expressed in an intensional language. We assume that the meaning or intension of (1) is the set of worlds in which pigs fly. We assume that this meaning is obtained from the meanings of 'pigs' and 'fly'. The simplest intensional semantics is to let the meaning of 'pigs' be the function ω which associates with each world the set of things which are pigs in that world and the meaning of 'fly' the function ζ which associates with each world the set of sets of things which fly in that world (the last complication is because (1) is a generic sentence and 'fly' is a property of a whole set of things).[6]

If it became true that pigs fly it would simply mean that we would be in a world which is different from the world we presently suppose ourselves to be in. It would be a world in which (1) is true, but not a world in which (1) means anything different. If however we associated with 'pigs' a <u>different</u> function, say the function η which picks out in each world the set of things which are birds in that world, then (1) in the very same world as at present would become true.

Extensional languages find this difficult to cope with. For an extensional interpretation cannot distinguish between a difference in extension caused by a change in facts and a difference in extension caused by a change in meaning. Many supporters of extensionality counter by denying that there is such a difference. The most notable defender of this view is Quine and I do not wish in this paper to take on the task of defending the distinction between the two ways of making pigs fly. It seems to me such an obvious distinction that I am almost at a loss to know how to argue for it. In any case, one who denies it certainly has the virtue of consistency and need not be interested in my ensuing discussion of how an advocate of the metalinguistic interpretation of propositional functors could incorporate the distinction.

In previous sections we considered a single intensional model with a set of possible worlds, and each world corresponded with a distinct model in the metalinguistic interpretation. $\mathfrak{M}$ was therfore a family of models corresponding to <u>one</u> intensional model. But when we make (1) true by altering the meaning of the words in it, we are considering moving to a <u>different</u> intensional model. So we ought, in the

metalinguistic case, to be thinking of a family $\mathfrak{Z}$ of classses of metalinguistic models. Many features of the members of $\mathfrak{Z}$ will need to be held constant. E.g. when we make (1) true by altering the meaning, we do not alter everything. The new model contains exactly the same set of worlds and, unless we have also changed the meanings of other worlds, the only change is a new function assigned to 'pigs'.

This enlargement allows us to prove a somewhat more general theorem of semantic generality, viz. a theorem to the effect that <u>any</u> collection of sentences of a propositional language may be regarded as a logic in that it may be characterized by a family of models. In section two it was mentioned that certain collections of sentences could be so treated. But not every collection of sentences can form a logic by the methods used there. For those methods required that a formula be valid iff true in every model in the collection (in the intensional model this means true of every world). This has the result that replacement of one theorem by another in any sentence cannot change its semantic properties and not every collection of sentences may respect this. It is in fact a rather strong condition for a logic and logics like the weaker Lewis modal systems S1-S3 do not satisfy it. What does give a general theorem is to take an intensional model and select one world as the 'designated' world of the model. One then defines validity in a class of models as truth in the designated world of each member of the class. It is not too hard to prove that any set of sentences in a propositional language can be characterized by a single intensional model with a designated world [5, p. 25]. This means (using the results of section two above) that any collection of sentences in a propositional language can be characterized by a class of metalinguistic models in which one model is designated.

More illuminating intensional characterizations can often be given by using a class of intensional models each with its designated world. In view of what was said in this section, such a characterization can also be achieved metalinguistically if we use a family $\mathfrak{Z}$ of classes of metalinguistic models where each class $\mathfrak{M}$ contains one of its members as the designated model.

5. DOUBLE INDEXING

In the last section we considered the possibility of expressing a class of intensional models by a family of classes of metalinguistic models. This had the effect of treating two rather different kinds of case as if they were similar; for the way the intensional semantics dealt with the two different ways of making sentence (1) true were very dissimilar. One was by change of worlds in the same model, the other was by

change of model. In this final section I want allude to the other way of dealing equally with the two cases, viz. to treat both cases in terms of a change of world. I shall do little more than make a few comments because it is an approach which has been advocated and used by Robert Stalnaker in dealing with several problems.

What I want to do in this section is to shew that Stalnaker's use in [16] and elsewhere of what has come to be called double indexing is equivalent to the classical intensional approach and therefore ultimately to Skyrms' completely metalinguistic approach. What I shall do first is set out the semantics (in a completely general form) of a doubly indexed propositional language and then shew how to express this semantics by one in which one of the indices is treated metalinguistically and the other intensionally. Actually we don't need to change the syntax or the intensional semantics of $\mathfrak{L}$ at all. All we need do is require that W be the set of all pairs taken from some set I. (I.e., $W = I \times I$.) Stalnaker takes I to be the set of possible worlds but an equally important use for them has been when they are moments of time. Logics with a semantics of this kind have been called two-dimensional modal logics [13]. We may therefore speak of two-dimensional semantics as well as double indexing.

Double indexing is used in Kamp's smenatics for 'now' [7] and more generally in the work of Vlach [17]. It has also been used by Åqvist and others [2]. These are all temporal uses with I understood as the set of all moments of time. A more properly modal use is needed for the semantics of 'actually'.[7]

However this use of double indexing is not our concern here. Our concern here is with the use of double indexing to represent metalinguistic facts, and the purpose of this section is to prove a theorem which is analogous to the theorem proved in section two, which will shew one way in which a two-dimensional semantics corresponds with a metalinguistic semantics.[8]

We are thinking here of classes of (one-dimensional) intensional models and want to investigate those metalinguistic properties which can be captured by a propositional functor in a two-dimensional semantics.

We suppose the language L_ω to be given as in section two.

We now suppose that $\mathfrak{M}_0$ is a collection of (one-dimensional) intensional models for L_0, indexed by I and each based on a set J of possible worlds. Although it seems reasonable to insist that J be the same in each model in $\mathfrak{M}_0$ there seems no reason to insist that $\mathfrak{M}_0$ be indexed by J. In fact it turns out that this does not matter, but it does make for a slight extra complication in the proof.

We shall first set out the class of models $\mathfrak{M}_{\omega}$ induced by $\mathfrak{M}_0$ and a collection $\mathfrak{R}$ of relations associated with the propositional functors. We associate with each k-place propositional functor δ a k+1-place relation R whose first term is an ordered pair of a model in $\mathfrak{M}_{n+1}$ and a member of J, and whose remaining terms are a k-tuple of classes of ordered pairs of a model in $\mathfrak{M}_n$ and a member of J. Where a is any formula of L_n we let

$$|a|_n = \{(M_n{}^i, j) : M_n{}^i \models_j a\}$$

We now shew how to define each $M_n{}^i{}_{+1}$ given $\mathfrak{M}_n$. (Remember that these models are one-dimensional intensional models.)

(i) If $a \in L_n$ then $M_n{}^i{}_{+1} \models_j a$

(ii) If a is $\delta a_1 \ldots a_k$, where $a_1, \ldots, a_k$ are all in L_n,

then where $j \in J$

$M_n{}^i{}_{+1} \models_j \delta a_1 \ldots a_k$ iff
$\langle M_n{}^i{}_{+1}, j\rangle \ R \ \langle |a_1|_n, \ldots, |a_k|_n\rangle$

What we must shew now is how a class of models of this kind leads to an equivalent two-dimensional model. The only reason why this result is not an immediate generalization of the result in section two is the fact mentioned earlier in this section that we cannot plausibly require that I = J. If I ≠ J then there are two possibilities regarding their cardinalities. Either $\bar{\bar{I}} \leq \bar{\bar{J}}$ or $\bar{\bar{J}} \leq \bar{\bar{I}}$. Suppose the former. Then let the two dimensional model be $\langle J \times J, V\rangle$ where V is defined as follows: Let π be a function from J onto I. Then

(i) If $a \in L_0$ and $j, j' \in J$ then

$$V(a) = \{\langle j, j'\rangle : M_0^{\pi(j)} \models_{j'} a\}$$

(ii) Let $\langle j, j'\rangle \in J \times J$ and let $a_1, \ldots, a_n$ be subsets of $J \times J$. Then
$\langle j, j'\rangle \in V(\delta)(a_1, \ldots, a_k)$ iff there exist $a_1, \ldots, a_k$ in some L_n such that $a_1, \ldots, a_k =$ $|a_1|_n, \ldots, |a_k|_n$

and $M_{n+1}^{\pi(j)} \models_{j'} \delta a_1 \ldots a_k$.

THEOREM 5.1a Where a is any wff of L_n then $M_{n+1}^{\pi(j)} \models_{j'} a$ iff $\langle J \times J, V\rangle \models_{\langle j, j'\rangle} a$.

The inductive proof of this theorem is exactly analogous to the proof of Theorem 2.2.

If $\bar{\bar{J}} \leq \bar{\bar{I}}$ then let the two-dimensional model be $\langle I \times I, V\rangle$ where V is defined as follows: Let π be a function from I onto J. Then

(i) If $a \in L_0$ and $i, i' \in I$

$$V(a) = \{\langle i, i'\rangle : M_0{}^i \vDash_{\pi(i')} a\}$$

(ii) Let $\langle i, i'\rangle \in I \times I$ and let $a_1, \ldots, a_k$ be subsets of $I \times I$, then

$$\langle i, i'\rangle \in V(\delta)(a_1, \ldots, a_k)$$

iff there exist $a_1, \ldots, a_k = |a_1|_n, \ldots, |a|_n$

and $\quad M_n{}^i{}_{+1} \vDash_{\pi(i')} \delta a_1 \ldots a_n$

THEOREM 5.1b Where a is any wff of L_n then $M_n{}^i{}_{+1} \vDash_{\pi(i')} a$ iff $\langle I \times I, V\rangle \vDash_{\langle i, i'\rangle} a$.

The proof of this theorem is also analogous to the proof of Theorem 2.2.

We can also obtain an analogue of Theorem 2.1, and in fact the situation in this direction is easier because we are already given a two-dimensional model in which both indices are taken from the same set. Given a model $\langle W, V\rangle$ in which $W = I \times I$ then we can define a family $\mathfrak{M}_0$ of metalinguistic one-dimensional intensional models in the following way (for $i, j \in I$)

(i) If $a \in L_0$ then

$$M_0{}^i \vDash_j a \text{ iff } \langle W, V\rangle \vDash_{\langle i, j\rangle} a$$

(ii) With each k-place functor δ we associate the k+1 relation R between models of L_{n+1} and k-tuples of classes of models of L_n defined as follows:

$\langle M_n{}^i{}_{+1}, j\rangle$ R $\langle A_1, \ldots, A_k\rangle$ iff there $a_1, \ldots, a_k$ of L_n such that $A_1, \ldots, A_k$ are respectively

$|a_1|_n, \ldots, |a_k|_n$ and
$\langle W, V\rangle \vDash_{\langle i, j\rangle} \delta a_1 \ldots a_n$.

THEOREM 5.2 $\langle W, V\rangle \vDash_{\langle i, j\rangle} a$ iff

$M_n{}^i \vDash_j a$

Obviously the results of section two can be applied to the one-dimensional models in order to reduce them completely to metalinguistic models.

What theorems 5.1 and 5.2 do is provide the conditions under which what can be said about the semantic properties of a two-dimensional functor can also be said metalinguistically. They also shew the form in which metalinguistic properties must be cast in order to describe them with a second index. One important fact which emerges is that the metalinguistic facts must be organized in a rather strict hierarchy. These equivalence results presuppose that we are accepting some kind of levels-of-language framework in the metalinguistic talk. This is one of the points Skyrms stresses in the section of his paper [14, pp. 382-387] when he discusses Montague's work on metalinguistic treatments of modality. The metalanguages used by Skyrms and used in this paper prohibit the kind of self-reference that Montague makes use of in [10]. So what has been shewn is that metalinguistic properties which are subject to these constraints are equivalent to intensional properties.

So much for the formal situation. The philosophical question is: how do we understand the second index? Or, in other words: what is it to associate a model with a world? One answer is a sort of pragmatic one (and indeed Stalnaker seems to think of his work as more properly described as pragmatics than semantics) to the effect that we are thinking of a speaker interpreting the words of $\mathfrak{L}$ in some particular way and the model associated with each world represents the way that speaker is using $\mathfrak{L}$ in that world. The reason, I suspect, that Stalnaker wants to call it pragmatics is that the _meaning_ of a word in any one world imposes no constraints on what it could mean in another world. It may be that we cannot be like Humpty Dumpty and use a word exactly as we please, but presumably how we do use it does not put limits on how we might use it. This certainly means that we are treating the two indices very differently because whether or not the sentence 'pigs fly' is true in another world does depend on how we are using the words in _this_ world and not in that other world.

Now Stalnaker has suggested in [15] that we might analyse statements about propositional attitudes as statements about what sentences mean. The idea seems to be this. Since all true mathematical statements have the same truth conditions (they are alike true in all possible worlds), and yet are not inter-substitutable in, say, belief contexts, then if the objects of beliefs are propositions construed as sets of worlds, we must find two different propositions, one of which

is associated with say $2 + 2 = 4$, and the other associated with, say, $\sqrt{144} = 12$. One appealing suggestion is simply that because those two sentences are distinct then there is a world in which one of them expresses a truth but not the other. This will not be a world in which these sentences mean what they do in our world, because in our world they both express the same proposition.

I have to admit that at this point I am not sure whether this _is_ the proper way to explain mathematical or logical beliefs but that is a topic in itself. The relevance to the present paper is simply that the generalization of the Skyrms result to the two-dimensional case can be used to shew at least one way in which an explicitly metalinguistic semantics and a two-dimensional semantics for propositional attitudes (when considered as involving a particular kind of propositional functor) can be shewn to be equivalent.

6. CONCLUSION

What do formal equivalence proofs of this kind discussed in this paper really shew? Well, I guess it is the moral alluded to earlier, that if one really wants to one can do things either way. That is not to say that one way may not be more convenient, more elegant, or more illuminating. But questions of this kind are best decided on the individual merits of each case. The general theorem licenses both the supporter of the intensional viewpoint and the supporter of the metalinguistic viewpoint to pursue the approach each favours, knowing that, if they so wished, each could translate the one approach into the other. Providing information of this kind is precisely the way logic is useful to philosophy.

NOTES

1. Skyrms does not in [14] consider Quine's "third grade" of modality in which quantifiers reach in to modal contexts, and neither will I. Skyrms has promised to deal with the third grade in a sequel. I am fairly sure that whatever devices make the third grade tractable for the modal operators will also carry over to the general case.

2. Skyrms uses the characterization of S5 found in Kripke's 1959 article [8] in which the worlds are actually thought of as models. However, the results, as shewn in the present paper, still hold whatever the worlds are.

3. Although this is correct, I have sacrificed some of the set-theoretical devices used in [5] to make the account more rigorous (in particular I have not bothered to insist explicitly that formulae be certain rather complicated sequences built up from the symbols). I hope the slightly more intuitive version presented here will be more easily apprehended without loss of precision.

4. In an intensional model a relation R between a world and a k-tuple of classes of worlds is used in [4, p. 348]. It is a variant of what are called 'neighbourhood semantics' [12, p. 13]. The equivalence between such an R and the more standard terminology of neighbourhood semantics is noted in [5, p. 27]. Note that in this proof we do not need to worry about what happens to those A's which are not $|\alpha|_n$ for any α. (This is similar to [4, p. 351] and [5, p. 21].) In [4, p. 364] an attempt is made to express metalinguistic facts intensionally, but no correspondence of the present kind is proved.

5. In [11, p. 160f] Scott suggests that <u>logical</u> <u>operators</u> should be those whose meaning should not depend on the structure of I.

6. A fuller discussion of questions of this kind within the framework of Montague Grammar is found in [3].

7. David Lewis in [9, p. 185f] suggests that 'actual' should be treated as Kamp treats 'now' in [7], though Lewis does not give an explicit formalization. Åqvist in [1, p. 10] defines an actuality operator which always goes back to a designated 'actual' world. This is interpreted two-dimensionally by Segerberg in [13, p. 77].

8. Some authors have in fact chosen to regard one of the

indices as models. Thus Crossley and Humberstone [6, p. 17f] introduce what they call global operators. The existence of papers like this is in part my motivation in proving the kinds of equivalences stated in Theorems 5.1-5.3.

REFERENCES

[1] Åqvist, L.E.G. son: 1973, 'Modal Logic with Subjunctive Conditionals and Dispositional Predicates', Journal of Philosophical Logic 2, 1-76.

[2] Åqvist, L.E.G. son: 1976, 'Formal Semantics for Verb Tenses as Analyzed by Reichenbach', in T.A. van Dijk (ed.), Pragmatics of Language and Literature, North Holland, Amsterdam, pp. 229-236.

[3] Carlson, G.N.: 1978, Reference to Kinds in English, Indiana University Linguistics Club, Bloomington, Indiana.

[4] Cresswell, M.J.: 1970, 'Classical Intensional Logics', Theoria 36, 347-372.

[5] Cresswell, M.J.: 1973, Logics and Languages, Methuen, London.

[6] Crossley, J.N. and I.L. Humberstone: 1977, 'The Logic of "Actually"', Reports on Mathematical Logic 8, 11-29.

[7] Kamp, J.A.W.: 1971, 'Formal Properties of "Now"', Theoria 37, 227-274.

[8] Kripke, S.A.: 1959, 'A Completeness Theorem in Modal Logic', Journal of Symbolic Logic 24, 1-14.

[9] Lewis, D.K.: 1970, 'Anselm and Actuality', Nous 4, 175-188.

[10] Montague, R.M.: 1963, 'Syntactical Treatments of Modality', Acta Philosophica Fennica: Modal and Many-Valued Logics, pp. 153-166.

[11] Scott, D.S.: 1970, 'Advice on Modal Logic', in K. Lambert (ed.), Philosophical Problems in Logic, Reidel, Dordrecht, pp. 143-173.

[12] Segerberg, K.: 1972, _An Essay on Classical Modal Logic_ 3 vols., University of Uppsala, Uppsala.

[13] Segerberg, K.: 1973, 'Two-Dimensional Modal Logic', _Journal of Philosophical Logic_ 2, 77-96.

[14] Skyrms, B.: 1978, 'An Immaculate Conception of Modality', _Journal of Philosophy_ 75, 368-387.

[15] Stalnaker, R.C.: 1976, 'Propositions', in A.F. Mackay and D. Merrill (eds.), _Issues in the Philosophy of Language_, Yale University Press, New Haven, pp. 79-91.

[16] Stalnaker, R.C.: 1978, 'Assertion', in P. Cole (ed.), _Syntax and Semantics_ Vol. 9, Academic Press, New York, pp. 315-332.

[17] Vlach, F.: 1973, _'Now' and 'Then': A Formal Study in the Logic of Tense Anaphora_, Ph.D. dissertation, University of California, Los Angeles.

Paul Gochet

THE SYNCATEGOREMATIC TREATMENT OF PREDICATES

In *Philosophy of Logic*, Quine approvingly describes a use of the *categorematic-syncategorematic* dichotomy in which the latter is equated with the distinction between expressions which are *names* and expressions which are *not names*. He then says that "What distinguishes a name is that it can stand coherently in the place of a variable in predication, and will yield true results when used to instantiate true universal quantifications."[1] That statement is followed immediately by the claim that "Predicates are not names". We can therefore safely avail ourselves of Küng's terminology and agree with him that "On the view of Quine and Goodman predicate signs can be regarded as syncategorematic".[2]

What does this view amount to? First, it amounts to a denial of the well-formedness of a quantification over predicates such as "Socrates is wise therefore Socrates is something". Second, it amounts to a denial of the claim that predicates *name* or *denote* sets or properties. As an alternative characterization of the semantic function of predicates, Quine writes: "Whereas a singular term purports to name an entity, abstract or concrete, a general term does not; but a general term is *true of* an entity or of each of many, or of none".[3]

If, however, general terms (common nouns or predicates) are *divorced* from sets and properties, and *connected directly* to the individuals they are true of, the following question arises, as Küng puts it: "Why is a certain predicate sign true for some individuals and not for others?"[4] In *Ontology and the Logistic Analysis of Language*, Küng has addressed himself to this problem and offered an original solution in the spirit of Wittgenstein's picture theory.

In this essay, I will re-examine Küng's solution and bring Goodman's analyses to bear on it. Later I will come to Quine's account of the functioning of predicates and relate it to his criterion of ontological commitment. I will defend the criterion against Vuillemin's objections but criticize it on other grounds. Finally I will try to locate the features in Quine's account of the semantics of predicates which aroused objections and sketch a defence and a program for a revision which meets Küng's demands.

I have stated the problem raised by the syncategorematic treatment of predicates. To solve it Küng sketches a new

B. K. Matilal and J. L. Shaw (eds.), Analytical Philosophy in Comparative Perspective, 61–80.

picture theory of predication. His theory is better understood if we bear in mind a few claims made by Wittgenstein in the Tractatus. The first claim worth recalling is about the composition of the atomic propositions. For Wittgenstein, atomic propositions are made out of words belonging to a single grammatical category, the category of names: "One name stands for one thing, another for another thing, and they are combined with one another. In this way the whole group like a tableau vivant presents a state of affairs (4.0311)".[5]

In a sequence of propositions, some names may recur. Far from being sheer redundancy, such a repetition of a name depicts a recurrence of the object named, as can be seen from this quotation: "Identity, I express by identity of sign, and not by using a sign for identity. Difference of objects, I express by difference of signs (5.53)".[6]

Küng modifies Wittgenstein's picture theory in two respects: first, he substitutes bicategorial semantics for the monocategorial semantics advocated by Wittgenstein for atomic sentences. Correlatively, he defines two concepts of identity: numerical identity and qualitative identity. Küng presents his account as a refinement of the picture theory: "...Wittgenstein, who, in our opinion, assumes ideally only one category of terms, viz. names of individuals, does not think of the relation of equality between signs as picturing something".[7] But, "if two categories of signs are distinguished, e.g. individual names and predicate signs, then the possibilities of representation are increased, and also become more complicated. In this case, the equality of signs of the one category will still indicate numerical identity, but the equality of the signs of the other category may then have the function of picturing something".[8] They will picture qualitative identity, i.e. recurrence of the same property as opposed to numerical recurrence of the same individual.

The following pair of sentences exemplifies the contrast between these two kinds of identity:

(1) Albert is intelligent, Bruno is intelligent.

(2) Albert is intelligent, Albert is courageous.

In sentence (2), the equality of the tokens of the name "Albert"--to use Küng's terminology--indicates the numerical identity of the referents of the token-words "Albert". In sentence (1), the equality of the tokens of the predicate "intelligent" pictures the qualitative identity or ressemblance of the concrete properties located "in" Albert and Bruno respectively.

What is a concrete property? Küng gives us a hint: "Remark how two equal concrete properties may be independent from one another: there are happenings which affect the

intelligence of Albert but not the intelligence of Bruno and vice versa".[9] The nature of these concrete properties remains to be explained. The notion of inherence which Küng expresses with the preposition "in" also needs to be clarified.

Küng's account is claimed to be "compatible with the rejection of Platonistic entities".[10] Moreover it is supposed to provide "what the contemporary Nominalists like Goodman fail to provide, namely an ontological justification for distinguishing true and false predication."[11] If one wishes to assess the respective merits of the two rival theories, that of Goodman and that of Küng, in a fair way, one should set them off correctly and not be misled by irrelevant differences.

Admittedly, Goodman advocates a syncategorematic account of predicates like Quine but, unlike Quine, he uses the word "predicate" to refer to copulae. The ordinary predicates, i.e. the descriptive ones, are relocated as names which, *qua* names, are, of course, not syncategorematic expressions, but, on the contrary, paradigms of categorematicity. Moreover Goodman advocates a very subtle ontology and combines it with a simple semantics. For instance, he recognizes the contrast between particulars and qualia, which roughly corresponds to Küng's opposition between individuals and properties, but he ascribes qualia as well as particulars to the Russellian type of individuals, a move which, presumably Küng would not be prepared to make.

Consequently, I cannot see what is gained by introducing as Küng does, the concept of *concrete property*. Goodman's *qualia* would account for the same semantic facts as Küng's concrete properties. Moreover Goodman's qualia seem to be superior to the latter in so far as their relation to particulars raises no problem, it is simply the *mereological relation* which is duly axiomatized in the calculus of Individuals, whereas the *inherence* relation between properties and individuals remains completely unexplained.

We have just seen that Goodman would paraphrase "Albert is intelligent" into "The quale intelligence is a part of the particular Albert". On this account, we claimed that the semantic function of the ordinary language predicate "intelligent" raises no problem: it is paraphrased into a noun which *denotes* an individual. No appeal has to be made to the metaphorical notion of picturing.

Admittedly, there are a few predicates left, those which I called "copulae", such as "is a part of" (<), "is a proper part of" (<<) and so on. All of them are definable in terms of the basic predicate "overlaps" (O) which is characterized by Goodman as "symmetric and reflexive but not transitive".[12] For instance, "<" is defined in this way

$$x < y = (z)\ (z \mathrm{O} x \supset z \mathrm{O} y)$$

It looks as if Goodman was immune to Küng's criticism of the descriptive predicates since they are no longer predicates in Goodman's treatment, but remained open to his criticism for the predicate "overlaps" and its derivatives. Later I will show that this objection can be met if we revive and adjust to our purpose a theory of predication outlined in 1949 by H. Leblanc, a theory which will prove congenial to some claims made by Küng himself when he is discussing the nature of the identity-relation.

Küng criticizes both Goodman and Quine for their _syncategorematic_ account of predicates. Hence a full response to Küng's objections against Goodman's theory of predication requires a detailed examination of Quine's treatment of predicates. That treatment is really the paradigm case of a syncategorematic treatment. One should here bear in mind that Quine developed his theory to its extreme limits: even class abstracts and relation abstracts are shown to be amenable to a syncategorematic treatment in certain cases. I have in mind here the _virtual_ theory of classes and relations which was presented by Quine in _Set Theory and Its Logic_ after a suggestion due to R. Martin.

The virtual theory of classes and relations emerged out of a cluster of doctrines among which one finds the critierion of ontological commitment. Quine's criterion of ontological commitment reads as follows: "In general, entities of a given sort are assumed by a theory if and only if some of them must be counted among the values of the variables in order that the statements affirmed in the theory be true".[13] If we combine this statement with the claim that only _names_ are replaceable by quantified variables, we have to conclude that _predicates_ are not equipped to carry ontological commitment. A problem arises in connection with the law of concretion:

$$y \in \{x : Fx\} \equiv Fy$$

On the right side of the equivalence sign we have a predicate and on the left we have a class abstract which occupies a name position. Shall we read a hidden class-abstract in the predicate "Fy" and say that predicates carry ontological commitments to sets after all, or shall we demote the class abstract to the status of a syncategorematic expression comparable to "Fy"?

Occam's razor settles the issue. We will say, Quine suggests, that "the whole combination "y ∈ {x : Fx}" reduces, by ... the _law of concretion_, to "Fy", so that there remains no hint of there being such a thing as the class {x : Fx}".[14] From this, by a very quick and productive move, Quine invented the _virtual theory of classes_: "Turning the tables and

speaking not of elimination but of introduction, we may view '∈' and class abstraction simply as fragments of a combination defined _in toto_ thus 'y ∈ {x : Fx}' for 'Fy'".[15]

Can we extend this to relation abstracts? Quine answers positively: "Just as the law of concretion for classes was used... to provide a joint definition of membership and class abstraction, so the law of concretion for relations provides a joint definition of the attribution and abstraction of relations: 'z {xy : Fxy} w' for 'Fzw'".[16] Before taking a stance on this question and deciding _with_ Quine that the virtual theory of relations is an innocuous extension of the virtual theory of classes or deciding, _against_ him, that it is a dubious step which requires independent justification, I will bring Vuillemin's argument to bear on the issue which goes to the heart of the matter.

Vuillemin disputes the claim that Quine's criterion of ontological commitment provides both necessary and sufficient conditions of an ontological commitment. According to Vuillemin, Quine's criterion supplies _sufficient_ conditions only. In other words, a statement may carry a commitment to more than the entities which have to be counted as values of the variables for the statement to be true. Take for instance the following statement in which a dyadic predicte occurs in attributive position:

(∃x) (∃y) (x is to the left of y)

Vuillemin agrees with Quine in saying that this statement carries an ontological commitment to individuals, but he contends that there is an _additional_ _commitment_ to the _order_ in which the individuals are located with respect to one another, and this additional commitment is not revealed, according to Vuillemin, by the criterion. If this is so the syncategorematic reading of the phrase "is a member of the relation from left to right" which explicitly refers to a relation allowed by the law of concretion is a pyrrhonic victory over relations since the syncategorematic reading of the dyadic _predicate_ "to the left of" itself is put into question.

We had better quote Vuillemin's argument here. Vuillemin wants to put forward as an argument "the whole philosophy of Russell" in so far as "it attributes to the use of relations, even in a first order logic, an ontological commitment...".[17] He sums up Russell's position which he adheres to in this passage:

> Consider the relation '_x_ is to the left of _y_'. Now suppose that all the ontology implied by our statement is absorbed by the individuals _x_ and _y_, as we...supposed was the case when we said that '_x_ and _y_ are red'. On this supposition the

> order in which the individuals _x_ and _y_ occur in the statement should be immaterial, barring the supposition that this order is involved in their nature as individuals... so that one does not escape a dilemma: either the relations are external, and ordering individuals which are themselves extraneous to that order, [they] introduce into the universe something that is not reducible to particulars. Or else they are internal, but it is then necessary to account for the asymmetric character of that order, by making it the property of each individual itself (monadism), or the property of the collectivity of individuals (monism); but in each case, one loses the sense of the order that is fundamental for the determination of asymmetry.[18]

The law of concretion enables us to express Vuillemin's insight in formal terms. Let us also note that the virtual theory of classes appears, paradoxically, to lend support to this attack on Quine's criterion. The law of concretion quoted above from Quine's _Set Theory and Its Logic_ reads as follows $z \{xy : Fxy\} w \equiv Fzw$ but the left sentence can also be written $<z, w> \in \{xy : Fxy\}$ and again, if we use Kuratowski's definition of the ordered pair: $\{\{z\}, \{z, w\}\} \in R$. But it should be remembered that the virtual theory of classes applies only to the class abstracts located on the right side of "$\in$", _to the exclusion of those located on the left_, the latter being irreducibly commital.

Is the occurrence of a class abstract _on the left_ of the sign "$\in$" a sufficient justification for the claim that polyadic predicates resist a syncategorematic account more than monadic predicates do? I do not think so. Vuillemin's argument shows that the _virtual theory of classes_ is insufficient to strip a sentence containing a relation abstract of its apparent ontological commitments, but it fails to establish that the _virtual theory of relations_ suffers from the same weakness. What cannot be achieved by the _virtual theory of classes_ may very well be achieved by the _virtual theory of relations_. The simple fact that relations _differ_ from classes, that they are more complex and have properties which classes do not have, such as asymmetry, for instance, is not enough to justify Vuillemin's move which, recast in Quine's terms, amounts to accepting the lesson of the concretion law in one case and to distrusting it in the other.

The _importance_ of the notion of _order_ cannot be disputed. Even to give the satisfaction conditions for monadic open sentences, i.e. monadic predicates, it is standard policy to use sequences, i.e. ordered n-tuples. But this _fact_ again does not suffice to establish that predication carries ontological

commitment to sequences. It merely proves that *metalinguistic* statements about satisfaction-conditions do. As Quine observes: "When I say that the pair <3,5> satisfies the sentence 'x < y', I am assuming for the time being that the sentence 'x < y' belongs to the object language and that the domain of objects of the object language includes the numbers 3 and 5, but I do not need to assume that this domain includes the pair <3,5>. The pair belongs to the apparatus of my study of the object-language, and that is enough".[19]

Let us sum up what has been achieved so far. We have agreed that Küng's dissatisfaction with the syncategorematic account of the functioning of predicates was justified. Küng's own solution has been examined next and compared with Goodman's. Finally the objection put forward by Vuillemin against the syncategorematic account of polyadic predicates has been answered. Shall we have to embrace the syncategorematic account of predicates after all, in spite of the misgivings evinced in the beginning? Not before we have scrutinized the credentials of Quine's account. As we shall see, they lie open to serious objections.

Quine subscribes to these three theses:

(1) Names only--together with expressions which play the same grammatical role as names such as class abstracts--are accessible to quantification.

(2) Predicates are not names.

(3) *All* quantified variables, and they only, can carry ontological commitment.

To thesis (1), one might object that the following sentence (a) makes perfectly good sense and conclude that the prohibition of quantification over the predicates is an unsupported dogma:

(a) Socrates is wise therefore Socrates is something.

There is, however, a reply. In *Word and Object*, Quine says: "For predication the verb may even be looked on as the fundamental form, in that it enters the predication without the auxiliary apparatus 'is' or 'is an'. The copula 'is' or 'is an' can accordingly be explained simply as a prefix serving to convert a general term from adjectival or substantival form to verbal form for predicative position. 'Sings', 'is singing' and 'is a singer' thus all emerge as verbs, and interchangeable ones apart from some subtleties of English idiom".[20] Now consider sentence (b)

(b) * Socrates talks therefore Socrates is something.

This sentence is clearly ill-formed. The substitution of "something" to verb-phrases generates ungrammatical sentences. This fact provides an argument to substantiate thesis (1) if we follow Quine in treating 'is wise' as a verb as he suggests in the above-mentioned statement. One might, however, object that this reply ignores the possibility of rephrasing the above-mentioned sentence into sentence (c) which is unquestionably well-formed:

(c) Socrates talks therefore Socrates does something.

Let us leave the question undecided for the time being and proceed to the examination of an objection recently raised against thesis (2). In "Ontological Burden of Grammatical Categories", Waragaī tries to break the link which Quine claims to exist between quantification and ontological commitment to entities. Waragaī contends that quantification over the prediates is both well formed and not commital, at least it does not commit us to <u>entities</u>.

Waragaī's position is summed up in this quotation: "If we replace a word designating some so-called <u>substantia prima</u>, e.g. Socrates, in a sentence containing this word, say 'Socrates is wise', with 'something' (<u>aliquid</u>, or better <u>aliqua res</u>), then we may be said to be commiteed to some entity by use of the sentence 'something is wise', but as to 'wise', the resulting sentence which we get by replacing this word with 'something' does not make us commit ourselves to any kind of entity... It only says that Socrates is in some mode of being. Only quantification of the word for the substantia prima forces us to commit ourselves to entities".[21]

Waragaī makes two separate claims: (1) quantifying over predicates does not commit you to <u>entities</u>, (2) it commits you to <u>modes</u> of being. To substantiate the former, one could observe that quantified variables substituted to a predicate such as "<u>Wise</u>" cannot occupy the <u>subject position</u> in a new predication, unlike quantified variables substituted to singular terms, be they concrete such as "<u>Socrates</u>" or abstract such as "<u>wisdom</u>".

From sentence (1), you can derive sentence (2)

(1) Socrates displays wisdom, wisdom is a virtue

(2) $(\exists \emptyset)$ (Socrates displays $\emptyset$ and $\emptyset$ is a virtue)

but nothing analogous can be done with "Socrates is wise". One can say "Being wise is being virtuous" and write the inference leading from sentence (3) to sentence (4)

> (3) Socrates is wise, being wise is being virtuous
>
> (4) $(\exists F)$ (Socrates is F, being F is being virtuous).

Notice, however the difference: "is a virtue" is predicated of $\emptyset$ in (2), but "is virtuous" is not predicated of "F" in (4). The second "is" in (3) and (4) is "$\supset$", not "$\in$". In formal terms (4) becomes (5)

> (5) $(\exists F)$ [Socrates is F, and $(\forall x)$ (x is F $\supset$ x is virtuous)]

Admittedly, Quine's criterion of ontological commitment does not explicitly contain the requirement that committal quantified variables be capable of occupying subject position, but I will argue that far from being an ad hoc requirement added for the sake of argument, that requirement can be given independent justification.

The connection between ontological commitment and subject position is formulated in Largeault's Enquête sur le nominalisme:

> It isn't the compression of elements into the collective unity of a set that is the platonistic feature of the principle of comprehension (for this unity could remain nominal, like that of an Ockhamite universal), but the fact that this unity is, in its turn, treated as an element of another class: the unit of the new class ceases entirely to be nominal at the moment when it is taken as an element. There is an analogy between this ontological phenomenon by which a variable, which is a place-marker so long as it is free, becomes purely designative from the moment that it is bound.[22]

To turn Largeault's suggestion into an argument, one should scrutinize the justification which can be given to support Quine's criterion of ontological commitment. The best justifiction I can think of can be found in Davidson's "The Method of Truth in Metaphysics". Davidson examines the ontoloigcal requirement of a homophonic theory of truth along Tarski's lines. If you want to give the truth-conditions for a finite language, he notes, you do not need any ontology. A

finite list of T-sentences of the form "'Socrates is wise' if and only if Socrates is wise" will suffice.

Again, a theory of truth which is ontologically free will be sufficient if you want to account for an infinite language where quantification does not enter into play. But where it does, ontology comes in:

> Multiple generality, however, the admission of whatever, in natural language, is treated by theory as sentences with mixed quantification and predictes of any degree of complexity, totally changes the picture. With complex quantificational structure, the theory must match up expressions with objects. But there is no need, as long as the underlying logic is assumed to be first order, to introduce entities to correspond to predicates.[23]

It is understood that the need for an ontology of sets or properties is felt when one reaches second order logic. But a sentence in which quantifiction over the predicates is allowed is not, for that matter, second order logic. You can quantify over the predicate, or better over common noun, i.e. general terms, in a first order logic built along the lines of Lesniewski's ontology. This is precisely the strategy adopted by Waragaī. I conclude that ontological commitment to classes or properties is not tied up with quantification alone, but with quantification associated with higher order predicates. On the ground of this additional requirement, I think we can agree with Waragaī that the sentence "Socrates is something" does not commit us to <u>entities</u> such as properties.

In some places Waragaī suggests that quantification over predicates commits us to <u>modes</u> <u>of</u> <u>being</u>, in other places he suggests more soberly that predicates <u>are</u> <u>related</u> <u>to</u> <u>their</u> <u>extension</u>. Let us quote his statement: "In general, names, i.e. categorematic expressions (in subject or predicate position), are related... to the reality as to their extensions, but those names which can be the subject of the sentence 'x ϵ y' are related to the reality as entity names".[24] This conclusion seems to run counter to Davidson's claim expressed in this passage: "...for large stretches of language anyway, variables, quantifiers, and singular terms must be construed as referential in function; not so for predicates".[25] Who is right, Waragaī or Davidson?

Davidson argues that as long as the underlying logic is first order, a Tarskian theory of truth <u>does</u> <u>not</u> <u>force</u> us <u>to</u> introduce entities as referents of the predicates. An ontology of individuals will suffice. <u>This</u> <u>is</u> <u>not</u> <u>enough</u>, <u>however</u>, <u>to</u> <u>prove</u> <u>that</u> <u>predicates</u> <u>do</u> <u>not</u> <u>refer</u>. Another view is compatible with Tarski's findings, namely the view that

predicates are referential but that they refer not to classes but to individuals. Martin who put forward this view claims that predicates are related to individuals by the semantic relation of "multiple denotation". According to Martin, "predicates are said to denote the objects to which they apply, e.g. 'dog' can be said to denote severally the particular dogs Marni, Randy, Fido, etc....".[26]

Martin's notion, however, is only terminologically different from Quine's talk of predicates "being true of individuals". Quine says that "a general term is true of each, severally, of any number of objects".[27] Now Quine alternatively says that general terms, though they do not refer, have an extension. In the light of these considerations I conclude that there is no substantial difference between Waragaï when he speaks of general terms related to the reality as to their extensions (versus to their referents) and Davidson when he says that predicates of first order languages do not refer. Much ado for nothing? Not quite, we have learned at least that Küng's talk about Quine's syncategorematic account of the functioning of predicates is misleading. On Quine's view, predicates do not refer but they have an extension at least and to that extent they are not fully syncategorematic in the sense in which the descriptor sign "ι" is fully syncategorematic in Russell's theory of description. But Quine's equating "categorematicity" with "namehood" and with "accessibility to quantification" invited the confusion. Martin's terminology clears it up. It can also be praised--if I may digress a little--for the very intuitive account it gives of the referential load of plural predicates such as "co-author" in the sentence "Russell is a co-author".

I have adopted Martin's account of the semantics of predicate, as far as descriptive predicates--as opposed to copulae--are concerned. I am well aware that Küng's puzzle can be reformulated against Martin's theory in this way: "Why do certain predicates severally denote some individuals and not others?" Here similarity and intension have to be brought to bear on the issue as Küng's iconic account of predication does not seem to have much explanatory power since it takes for granted similarity in rebus. Küng's revival of the picture theory offers no explanation to the formation of natural kinds. It presupposes them.

There is a special class of predicates which traditionally have been studied in logic and ontology and still are and rightly so. I have in mind "∈", "⊂", '='. I will call the copulae. For these predicates, I am going to offer a syncategorematic theory which is a revival and an extension of a theory put forward by Leblanc. The latter discusses "=" and "⊂" only. I will try to extend his theory to "∈". As descriptive predicates such as "walks" are reckoned as partly syncategorematic in so far as they embody a copula such as "="

or "$\in$", that account does justice to Quine's insights to a certain extent.

Let us consider sentences (1)-(4) where X and Y deonte classes

(1) $\underline{X} \subset \underline{X} \cup \underline{Y}$
(2) $\underline{X} \subset \underline{Y}$
(3) $\underline{X} = \underline{X}$
(4) $\underline{X} = \underline{Y}$

Sentence (1) serves to state that *set* $\underline{X}$ is contained in *set* $\underline{X} \cup \underline{Y}$. But as the simple *sign* "$\underline{X}$" was already contained in the complex *sign* "$\underline{X} \cup \underline{Y}$", Leblanc says that the inclusion sign is here simply used to *re-assert* the inclusion of $\underline{X}$ in $\underline{X} \cup \underline{Y}$ which was somehow *pictured* by the inclusion of the *sign* "$\underline{X}$" in the sign "$\underline{X} \cup \underline{Y}$". In sentence (2), set $\underline{X}$ is described as included in set $\underline{Y}$, but in this case the sign "$\underline{X}$" was not contained in the sign "$\underline{Y}$". Hence the copula here plays a different role. According to Leblanc, it is used to *annul* some properties of the "$\underline{X}$" and "$\underline{Y}$" which *hide*, instead of picturing, the inclusion of one referent in the other. The same contrast can be made to account for the use of the predicate "=" in the sentences (3) and (4).

Predicates, which we would better call "copulae", are conceived by Leblanc as devices needed to *restore* between language and the world the *isomorphism* advocated by the picture theory of meaning: "To sum up," Leblanc writes, "... our symbols being already too expressive as concrete entities must be governed by non-designative devices like predicates which ensure a closer correspondence between notation and reality. Predicates do not name properties or components of individuals; they are linguistic tools illuminating a basically opaque notation, correcting its physical *defects and strengthening its* haphazard notational virtualities."[28]

As it stands, this formulation is rather *misleading*. Leblanc says that in sentence (1) the copula *reasserts* the inclusion of the referent of the sign "$\underline{X}$" in the referent of the sign "$\underline{X} \cup \underline{Y}$". One can *re-assert* only what has already been *asserted* a first time. But the inclusion of the sign "$\underline{X}$" in the sign "$\underline{X} \cup \underline{Y}$" *does not assert* anything, though it *depicts* the inclusion somehow. Again in sentence (3), the predicate "=" could serve to re-assert the identity of the refernts of "X" only if that identity had already been *asserted* by the very repetition of the same sign "X". But a mere repetition of a sign cannot assert, it can only picture or express the identity of the referent.

One might also balk at the idea that the predicate "=" or the inclusion sign "$\subset$" can serve two opposite purposes: *re-assert* what is already expressed by other features or *annul* these features. On a closer examination, however, one

discovers here an insight which is worth preserving. I shall try to bring it out.

In sentence "*X* = *Y*" or "X $\subset$ Y", the predicate or copula is used, Leblanc said, to *annul* an unwanted feature of our symbolism. Which feature? In the first sentence the fact that the signs "*X*" and "*Y*" are different from one another. To annul a difference, amounts to restoring an identity. Compare sentences (1) and (2)

(1) Tullius = Tullius

(2) Tullius = Cicero

The former is a truth of logic, the latter is not. The latter is informative, the former is not. These logical and epistemological differences not withstanding, (1) and (2) display a peculiarity: unless the *two* terms related by the predicate stand for a *single* entity in the outside world, the sentence is not true. This peculiarity sets the identity relation apart from the other relations such as, for instance, the relation of brotherhood. G. Küng notices it and writes: "A reflexive relation... is not a real relation, but rather a relation of reason based on the conceptual duplication of an entity that in reality is a single entity."[29] I will argue that here we have the *defining characteristics of copulae*: they express a *conceptual duplication of what in reality is a single entity*. Unlike Küng, however, I do not link this feature with reflexivity. The equivalence relation expressed by "compatriot of" is a *real* relation, which can obtain between distinct individuals and yet it is a *reflexive* relation.

My account of the role of copulae can be traced back to Aristotle who writes: "As for 'being' *qua* truth and 'not being' *qua* falsity... they depend upon combination and separation. By 'combining' and 'separating in thought', I mean thinking them not as a succession but as a unity..." and below "But... the combination and separation exists in thought and not in things."[30]

The account of predication as *conceptual* or *linguistic duplication* can be applied to Goodman's predicate "is a part of" (<) very easily. The mereological predicate "is a part of" when inserted between two names gives rise to a true sentence if and only if the referent of the first name is a part of, i.e., *partly coincides*, with the referent of the second name. This is another case of a linguistic duplication which does not coincide with a case of duality in *rebus*. An analogous comment can be made about the sign for class-inclusion. Class inclusion is a sort of *partial identity*, as can be seen from the law $A = B.\equiv.[(A \subset B)$ and $(B \subset A)]$.

One might object that what I called "real relations" sometimes displays the same peculiarities as "conceptual

relations." Think of (1) "Henri is a compatriot of himself" or of (2) "England is smaller than Great Britain". My reply to the first objection is that reflexivity is not a sufficient condition for a predicate to enjoy the status of a copula. As to the sentence (2), it displays a mixture of real relation (smaller than) with a conceptual relation (is part of).

The class membership predicate "$\in$", which is not reflexive, apparently raises more problems. According to Bergmann there is an exemplification tie which serves to relate universals to particulars. Those who favor that theory might be tempted to say that "$\in$" denotes an exemplification tie in the way "brother of" denotes a blood tie. I will counter this by adducing evidence for the claim that "$\in$" is more like "=".

Consider the following sentences:

(1) Phosphorus $\in$ {x:x=Phosphorus V...V x = Pluto}

(2) Phosphorus $\in$ {x:x=Hesperus V...V x = Pluto}

(3) Phosphorus $\in$ {x:Planet x}

In (1) there is a duplication of the name "Phosphorus". In (2), there is a duplication again, although we no longer face two tokens of the same type word, but tokens of two types. In (3) there seems to be no duplication at all. Yet if we take up Martin's theory of multiple denotation, we can look at "Planet" as shorthand for "x = Hesperus V... x = Pluto". In other words, Martin's concept of multiple denotation contributes to bridging the gap between (2) and (3).

Differences of meaning left aside, these three sentences are also alike in truth-value. If it could be established that they also state the same fact, we would be justified in holding that the sign for class-membership plays the same role in each. My argument will collapse, however, if it can be shown that there is no such thing as identity of fact over and above identity in truth-value. There is in Church's Introduction to Mathematical Logic an argument meant to show just that. Church observes that one can go from "Sir Walter Scott is the author of Waverley" to "The number of counties in Utah is twenty-nine" simply by replacing denoting expressions by other denoting expressions which are co-extensive, i.e., which are co-referential, and by moving expressions around. One intermediary link will be the statement "The number, such that Sir Walter Scott is the man who wrote that many Waverly Novels altogether, is twenty-nine".[31]

In order to block Church's slippery slope argument just mentioned, I believe it is necessary and sufficient to add to the requirement of termwise coextensionality the further

condition of isomorphic coextensionality. The first and the last sentence quoted from Church are synonymous and termwise coextensive but they are not isomorphic.[32]

The notion of extensional isomorphism I borrow from Mleziva's "Problem fakt v logicke semantice."[33] Roughly speaking, two sentences are said to exemplify extensional isomorphism when they are co-extensive (bear the same truth-value) and when their parts are co-extensive. From this I draw the following definition of fact: two sentences express the same fact when they are related to each other by the relation of extensional isomorphism which is less than synonymy but more than material equivalence. One might object that (1) and (2) are not isomorphic, that they do not have quite the same logical form and that (3) is still more different. If, however, we take up Quine's virtual theory of classes and relations, the objection vanishes. From the point of view of ontology all the three share the logical form "F(a)" or, on a finer analysis, $(\exists y)(Fy \ \& \ (x)(Ax \equiv x=y))$.

The account of predication I am proposing can be described as a vindication of the syncategorematic account of the functioning of predicates where "predicate" is used to designate copulae. On my nominalist account, which combines Leblanc's views with Küng's, the predicates "=", "⊂", "∈" do not designate, they are part and parcel of the machinery of predication. As to predication, I described it as a conceptual duplication. In the predicate calculus, the duplication shows itself in the repetition of co-occurrent individual variables. The three examples below bring that out:

(1) Some men are wise
$(\exists x) \quad (\text{Man } \underline{x} \ \& \ \text{Wise } \underline{x})$

(2) All men are rational
$(x) \quad (\text{Man } \underline{x} \supset \text{Rational } \underline{x})$

(3) Socrates is wise
$(\exists y) \quad [\text{Wise } \underline{y} \ \& \ (x) \ (\text{Socratizes } x .\equiv. \ x = \underline{y})]$

A look at sentence (4) reveals another repetition, i.e., the repetiton of the predicate "Wise".

(4) Socrates and Plato are wise
$(\exists x)(\exists y)(\underline{\text{Wise}} \ x \ \& \ \underline{\text{Wise}} \ y \ \&$
$(v)(\text{Socratizes } v \supset v=x) \& (u)(\text{Platonizes } u \supset u=y)$

One should be allowed to say that the latter repetition also depicts something, but for the reasons given in the body of this essay, one should refrain from concluding that the repetition depicts some thing. On the negative side I thus agree

with Küng and Waragai, but not on the positive side. I fail to see any explanatory power in the notions of "concrete property" or "modes of being". To the repetition of the predicate "Wise" one could just as well associate the identity of the individuals taken severally, which those predicates are true of or multiply denote along the lines of Martín. The statement that individuals are multiply denoted by the same predicate because they share the same concrete property or mode of being may be true but as we have no independent access to properties such a statement fails to count as a solution to Küng's puzzle.

A statement such as "There are no unicorns" is no counter-example to my thesis. As Quine says in "On Not Learning to Quantify", "Truly unrestricted quantification is rare outside logic."[34] The above statement can be read as "a categorical with tacit first term." The categorical I have in mind is

Whatever exists is a non-unicorn

$$(x)\ [(\exists y)\ (y=x) \supset -Ux]$$

As a quantification theory requires a non empty universe, we will have at least one individual standing as a value for the variable "x" which occurs twice and serves to carry out the subject-predicate duplication.

It might be thought that I changed the topic when I moved from Küng's picture theory of predicate to Leblanc's picture theory of predication. I would like to show how the two theories are related. With this I will bring my discussion to a close.

For quite a long time Quine held the view that the ontology needed by physics was made of "physical objects, classes of them, classes in turn of the elements of this combined domain, and so on up."[35] A fact about microphysics led him to change his mind. The fact is reported in The Scientific American (1973) by Post.

If we consider two electrons x and y and two boxes East Box and West Box, we think of four possible locations for the electrons. Yet statistical findings, says Post, show that two of the four possibilities acknowledged by common sense collapse into one. From this Quine concludes that "we should think not of individual electrons x and y but of states of the boxes: ...states of being singly or doubly affected."[36]

This does not boil down to the replacing of an ontology of electrons by an ontology of states. States are predicates which are attributed to space-time regions, i.e. to sets of space-time points. But, as Quine observes, "we can invoke Cartesian coordinates and identify each space-time point with a mere quadruple of real numbers."[37]

Predicates that formerly attributed states to points or regions will now apply to quadruples of numbers, or to sets of

quadruples. As predicates are *syncategorematic* expressions which carry no ontological commitment, Quine is driven to a pythagorean ontology or even--if numbers are reduced to sets without ground elements--to "Hyperpythagoreanism" as he calls it. There is, however, another conclusion to draw. If we have the courage to accept Quine's "brave new ontology", it follows that Küng's problem is not so much *solved* as *dissolved*. I will end this essay by establishing this point.

Once it is agreed that the basic subjects of our predications are extensional entities par excellence, viz. numbers, it becomes queer to ask for a *fundamentum in re* for the physical properties which are attributed to them. Being the number of the planets or the number of the Muses can still be predicated of the number 9 and being a place where the temperature in degrees Kelvin is such and such can still be predicated of the quadruples of real numbers $\langle a_1, a_2, a_3, a_4 \rangle$ and $\langle b_1, b_2, b_3, b_4 \rangle$ but it would be *senseless* to look for something *in* the numbers themselves which makes them the bearers of physical properties. Küng's search for a justifiction of our predicating a predicate of some individuals and not of others is reduced to absurdity if it is understood as an attempt at isolating *in* the subjects some inherent property which makes the predication true or whose absence makes it false. In Quine's latest ontology, Küng's question no longer arises. This is not meant, however, as an argument in favor of Quine's new ontology but as a clue which reveals what hinges on the *syncategorematic theory of predicates*.

Quine realized the devastating consequences of Hyperpythagoreanism for metaphysics and tried to find an escape. Taking advantage of the distinction between *ontology* (the values of the bound variables) and *ideology* (the set of predicate constants enumerated in the lexicon of the theory), he suggested that we might *redefine metaphysics*: "We might consider... that the lexicon of natural science, not the ontology, is where the metaphysical action is."[38] The new task for metaphysics consists of the attempt at reducing the lexicon in accordance with the requirements of physicalism.

Such a *way out* will not do, however, if the account of predication defended in this paper is accepted. If predication is understood as a kind of duplication, the sharp contrast between *terms* and *predicates* has to be abandoned, as far as the descriptive predicates (versus copulae) are concerned. Predicates are not *fully* syncategorematic in the way the descriptor "ι" is. Hence the sharp contrast between ontology and ideology which underlies Quine's new definition of metaphysics has to be mitigated, or his austere hyperpythagoreanism has to be rejected.

University of Liege

NOTES

1. W.V.O. Quine: 1970, Philosophy of Logic, Prentice-Hall, Englewood Cliffs, N.J., p. 27.

2. G. Küng: 1967, Ontology and the Logistic Analysis of Language, E.C.M. Mays, tr., D. Reidel, Dordrecht, p. 155.

3. W.V.O. Quine: 1961, From a Logical Point of View, Harper Torchbooks, New York, p. 21.

4. G. Küng, op. cit., p. 156.

5. Wittgenstein: 1961, Tractatus Logico-Philosophicus, D.F. Pears and B.F. McGuinness, trs., Routledge and Kegan Paul, London, p. 43.

6. Wittgenstein, op. cit., p. 53.

7. G. Küng, op. cit., p. 164.

8. G. Küng, ibid., p. 164.

9. G. Küng, op. cit., p. 166.

10. G. Küng, ibid., p. 166.

11. G. Küng, ibid., p. 166.

12. N. Goodman: 1966, The Structure of Appearance, Bobbs-Merrill, New York, p. 47.

13. W.V.O. Quine, From a Logical Point of View, p. 103.

14. W.V.O. Quine: 1969, Set Theory and its Logic, Belknap Press and Harvard University, p. 16.

15. W.V.O. Quine, ibid., p. 16.

16. W.V.O. Quine, ibid., p. 21.

17. J. Vuillemin, La logique et le monde sensible, Flammarion, Paris, p. 43.

18. J. Vuillemin, ibid., p. 46.

19. W.V.O. Quine: 1970, The Philosophy of Logic, Prentice-Hall, Englewood Cliffs, N.J., p. 37.

20. W.V.O. Quine: 1980, Word and Object, Wiley, New York, pp. 96-97.

21. T. Waragaĩ: 1979, 'Ontological Burden of Grammatical Categories', Annals of Japanese Association for Philosophy of Science, 5, no. 4, 185-205.

22. J. Largeault, Enquête sur le nominalisme, Nauwelaerts, Louvain, p. 383.

23. D. Davidson: 1979, 'The Method of Truth in Metaphysics', in Contemporary Perspectives in the Philosophy of Language, P. French, T. Uehling, H. Wettstein, eds., University of Minnesota Press, p. 301.

24. T. Waragaĩ, op. cit., p. 47.

25. D. Davidson, op. cit., p. 301.

26. R. Martin: 1958, Truth and Denotation, Routledge and Kegan Paul, London, pp. 99-100.

27. W.V.O. Quine: 1960, Word and Object, Wiley, New York, pp. 90-91.

28. H. Leblanc: 1949, 'The Semiotic Function of Predicates', The Journal of Philosophy 46, 842.

29. G. Küng, op. cit., p. 165.

30. Aristotle, Metaphysics 1027b20, 1027b30, Loeb Classical Library, transl. by H. Fredernick, Harvard University Press, London, 1968.

31. A. Church: 1956, Introduction to Mathematical Logic, Princeton University Press, Princeton, p. 2.

32. P. Gochet: 1980, Outline of a Nominalist Theory of Propositions, D. Reidel, Dordrecht, p. 137.

33. Mleziva: 1969, "Problem fakt u Semantice", Theorie a Metoda, Prague, p. 74.

34. W.V.O. Quine: 1979, "On Not Learning to Quantify", The Journal of Philosophy, p. 429.

35. W.V.O. Quine: 1966, "The Scope and Language of Science", Ways of Paradox and Other Essays, Random House, New York, p. 231.

36. W.V.O. Quine: 1976, "Whither Physical Objects", in Essays in Memory of Imre Lakatos, ed. R.S. Cohen and M.W. Wartofwsky, D. Reidel, Dordrecht, p. 499.

37. Quine, ibid., p. 501.

38. W.V.O. Quine: 1979, "Facts of the Matter", Essays on the Philosophy of Quine, ed. R. Shahan and K. Merrill, Harvester, p. 160.

W. Lycan

THE PARADOX OF NAMING

There is overwhelming evidence that proper names must have senses or connotations that somehow contain contingent information about their referents. There is also overwhelming evidence that proper names cannot possibly have such senses or connotations. That is the paradox of naming.

The former, familiar evidence comes from Frege and Russell (and from their contemporary inheritors[1]) in the form of appeals to substitutivity failure, negative existentials, the informativeness of identity statements, and so on. The latter evidence comes largely from Saul Kripke,[2] who has defended the opposing Millian tradition by pointing out that if names abbreviated definite descriptions or otherwise expressed contingent properties of their referents, (a) certain sentences would be tautologous or contradictory that are not in fact tautologous or contradictory; (b) certain names that denote distinct individuals would denote the same individual; (c) serious problems would flow from the then presumed *equivocity* of names; and (d) a name would denote different individuals in different possible worlds.[3] Rather, Kripke maintains (1972), names are *rigid* designators; a name denotes an individual quite independently of any contingent property the individual might have, and hence denotes that same individual in any world in which it exists. Mill seems to have been right in contending that the semantic function of a name is simply to pick out its bearer, and it follows that names are rigid in the sense defined.

I shall not here rehearse any of the well known arguments on either side, but will try to sharpen the paradox in my own way.

1.

Suppose that on occasion *O*, *S* tokens 'Cicero was bald, but Tully was not,' in a suitably sincere tone and goes on to behave accordingly. Observing this, we would all naturally want to infer (A):

> (A) *S* believes that Cicero was bald and Tully was not.

What possible positions might we take regarding the semantics

B. K. Matilal and J. L. Shaw (eds.), Analytical Philosophy in Comparative Perspective, 81–102.

(and resulting truth-value) of (A)? Let's assume for convenience that 'Cicero' and 'Tully' have the same semantical status, whatever that status might be; imagine an appropriate story leading up to O.

(I) 'Cicero' and 'Tully' are connotationless, "Millian" names. (As I've said, it's clear that all Millian names are rigid designators in Kripke's sense, but the converse does not hold. A Millian name is one whose sole semantical function is to pick out its bearer; contrast 'the greatest prime smaller than 3,087,' which due to the necessity of arithmetic truths is rigid, but not Millian.)

(a) S doesn't really believe what (A) says he believes; (A) is false.

(b) S does really believe it; (A) is true and S accepts a contradiction,

(i) Despite appearances, S is irrational.

(ii) Sometimes belief in contradictions doesn't reflect badly on one's rationality.

(II) 'Cicero' and 'Tully' are not Millian names; i.e., they make (distinct) semantical contributions of some specific kind, over and above denoting their common bearer.

(a) They abbreviate flaccid (nonrigid) definite descriptions in (A), and perhaps also for S on O.

(b) They abbreviate world-indexed or otherwise rigidified descriptions, and thus are rigid though not Millian. (This is Alvin Plantinga's "Boethian Compromise," more on which below.)

(c) . . . ?

Opinion (I) is hard to maintain in the face of Russellian arguments. Indeed, Kripke himself has recently impugned it in propounding his "Puzzle about Belief."[4] I think we can take (I) a good deal farther than many people suppose,[5] especially if we take suboption (b-ii) seriously (I shall be returning to (b-ii)), but it is not intuitively a very appealing view and it has some very nasty consequences (see Baker (in press)). Our only very compelling motivation for (I) is our desire to avoid (II-a), the Russellian option ostensibly discredited by Kripke's original criticisms. What we really want is some

tertium quid, some hybrid or conciliatory synthesis somewhere and somehow in between Millianism and Russellianism. But what form would it take?

I propose the following constraint, which will help to bring out the difficulty of our dilemma.

> THE RULE OF EFFABILITY: Any semantical account of names that constitutes a tertium quid of the sort we are seeking must be representable as the assignment of a logical form to (A), which logical form is expressed in standard logic plus a univocal belief operator. [Qualifications: (1) No ontological holds are barred; logical forms may involve quantification over any entities one likes. (2) Logical representations need bear no discoverable grammatical relations to (A).]

My justification for this rule is substantially just that it's hard to think of any alternative format for a semantical theory. (Notice that the Rule does not forbid the use of intensional idioms; intensionality will come in, or may come in, by way of the introduction of intensional entities such as properties, propositions, possible worlds, or world lines. There is no fancy logic, however outré, whose semantic interpretation can't be written in standard logic.) Now, the paradox of naming is generated by the assumption that our only possible options anent sentences like (A) are (I), (II-a), or something effectively equivalent to (II-a)--in short, the assumption that any name that is not Millian in effect abbreviates a description. I am inclined to think that what gives this (usually unexamined) assumption its plausibility is the Rule of Effability or something like it. If a name is not a semantically simple, unstructured individual constant at the level of logical form, then it must be some combination of constants, variables and predicates; the only type of combination that seems at all appropriate is a Russellized description.

A word about "abbreviation": One excellent reason for denying hotly that names abbreviate descriptions is that it's fanciful at best to suppose that when an actual speaker tokens a sentence such as (A), an actual, psychologically real abbreviative operation has occurred within that speaker; surely nowhere in the speaker's mind has some Russellian quantificational structure triggered an abbreviating transformation that deletes that fairly rich structure and substitutes the tag 'Cicero' or whatever. (On this robust account, different quantificational structures within different speakers, or even within the same speaker at different times, would just happen to be scrunched up and lexicalized into the same short word.) So it may seem obvious that if this is what is meant by

'abbreviate' in (II-a), there must be a more plausible version of option (II)--one which takes 'abbreviate' less literally---and so there's no paradox yet. Quite so. But Kripke's objections to option (II-a) do not turn on the issue of the psychological reality of abbreviation; they're directed against even the much weaker suggestion that a proper name in an ordinary speech context is equivalent in _any_ semantical way to any particular flaccid description or body of contigently descriptive material. So let's hereafter understand (II-a) in this weaker, nonrealistic sense of 'abbreviate.' I shall return to the issue of abbreviation in section 6 below.

2.

Alvin Plantinga's "Boethian Compromise" (1978) actually provides a clever and ingenious _tertium quid_, (II-b), that satisfies the Rule of Effability. It capitalizes on the previously neglected distinction between being Millian and merely being rigid. That distinction had always been illustrated mainly by throwaway examples from arithmetic that turn on the (alleged) necessity of all arithmetic truth, and so it never seemed particularly important. What Plantinga noticed is that for any contingent identifying property that any ordinary thing has, there is a corresponding "world-indexed" property that is an essential property of that thing, and for each such world-indexed property there will be an identifying description of the thing that is rigid but (because it has descriptive content and refers only by exploiting that content) not Millian. (Example: To 'the winner of the 1962 Outboard-Motor-Eating Contest,' a flaccid description, there corresponds 'the winner of the 1962 Outboard-Motor-Eating Contest _at @_' (@ being our actual world), which is rigid; at a given world _w_, it picks out, not whatever denizen of _w_ wins the contest at _w_, but Attila Schwartzenegger, the man who actually won the contest here at @.) Thus Plantinga seems to slip neatly between the horns of our dilemma.

Unfortunately, as has not been generally noticed, Plantinga is as firmly committed to the contradictoriness of our friend _S_'s belief as is the Millian. Suppose 'Cicero' and 'Tully' abbreviate distinct world-indexed descriptions, both satisfied by their common referent. Both descriptions are rigid; that's what world-indexing is for. But if both descriptions are rigid, then _S_'s utterance on _Q_ is true in no possible world. It refers doubly to one and the same individual at each world and predicates contradictory properties of him. The proponent of (II-b) has just the same options _vis-á-vis_ _S_'s rationality as has the proponent of (I-b). In this sense, the mere rigidity of names is just as bad as full-scale Millianness.[6] The same point can be brought out by a slight reformulation of our original dilemma:[7] As before, (A) seems

to be true. But if names are rigid, then (B) _S_'s sentence is false at any possible world, and to believe what it expresses is to believe an explicit contradiction. Yet we may suppose that _S_ is fully rational and would not accept any explicit contradictions; (C) _S_ is committing no _fallacy_ or crass logical error, however ignorant he may be.

3.

Of course, there are plenty of different kinds of semantical theories all directed towards problems of referential opacity and related matters. So it may seem surprising that anyone should think we had exhausted all our options so soon. What about inscriptionalism? Frege's own shift-of-reference view? Hintikka's theory of world-lines? All the solutions that Quine considers in his various writings on this topic? Etc. Some of these, such as näive inscriptionalism, are plainly inadequate from the beginning. More commonly, though, a view will be seen to _collapse into the Russellian option_ even though the view seems more elaborate and mobilizes sophisticated apparatus. One such view is Frege's, as I read him. He hypothesizes that inside a (single) belief operator, a name shifts its reference from its customary bearer to its customary _sense_, an abstract constituent of the "thought" or proposition named by the complement clause as a whole; what causes substitutivity failure is that distinct names normally have distinct customary senses even though they may happen to have the same customary referents. But in Frege's examples,[8] the senses of names are expressed in the form of definite descriptions, implying that names _have_ the senses of descriptions. Thus Frege seems committed to (II-a) and open to Kripke's objections.[9]

A more interesting example of a sophisticated theory that seems to collapse into Russellianism is Hintikka's classic possible-worlds version of Frege's approach ((1969) and elsewhere). Hintikka holds that inside belief contexts names name, not the ordinary inhabitants of our world, but "individuating functions" or world-lines. Each world-line, corresponding roughly (but not exactly) to a Fregean individual concept, is a function from worlds to individual inhabitants of those worlds. (On Hintikka's view, actually, it is the world-line itself that corresponds to a real individual, since real individuals such as you and I come complete with counterfactual properties as well as the actual properties we have here at @. This leaves us with the uncomfortable question of what sorts of things the values of world-lines are. Hintikka's followers sometimes call them "individual _manifestations at_" worlds; or we might call them world-_slices_ of individuals.) Each world line, associated with a proper name N, is determined or generated by a particular "method of

cross-identification" or reidentification procedure governing the correct use of N (i.e., with a rule for answering questions of the form, ⌜Which denizen of w is (identical with) N?⌝). Different names of the same person can be associated with different methods of cross-identification in the minds of believers; the procedure for recognizing Clark Kent in a given world, in the mind of a believer who is unaware that Kent is Superman, is obviously very different from the same believer's procedure for recognizing Superman at a given world. The reason that names don't substitute in belief contexts is that in those contexts they name, not the same earthly person, but distinct (because diverging) world-lines.

Notice that by virtue of his appeal to the reidentification procedures speakers actually use, Hintikka is committed to the contingency of identity statements involving names and hence to the flaccidity of names, at least in belief contexts. (He remarks that "...even an allegedly rigid designator can be bent by epistemic considerations" (1975), p. xi). World-lines are generated by sets of reidentification criteria that human believers actually use in tracing other things and people, or so we are left to suppose. But any such criterion must be substantive and useful as an epistemic tool; as Russell kept saying in "The Philosophy of Logical Atomism," we don't recognize our friends by getting a look at their metaphysical insides and ogling their haeccities. We spot them by their appearances, their actions, their personalities, and so on--almost invariably, by clearly contingent properties they have. Thus, on Hintikka's view any name will apply at world w to whomever or whatever has such-and-such contingent properties at that world, rather than to whomever or whatever is intuitively the individual that the name denotes here at @. And this is just a version of thesis (II-a), the Russellian option; Kripke's original objections apply.[10]

It's all very discouraging.[11] But I think there is a genuine way out, one that I have defended obliquely elsewhere ((1981b); cf. (1981c)). It will take us at least briefly into philosophy of mind, because I think any real solution to the paradox of naming and/or to Kripke's "Puzzle" requires some attention to psychological matters.[12]

4.

Wilfrid Sellars and a number of subsequent authors have defended the view that the sentential complement of a belief-ascription serves as a sort of exemplar or sample of what is said to be believed (Sellars (1956), (1963), (1967), (1969), (1973); see also Davidson (1968) and Hill (1976)). For example,

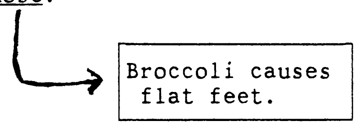
(1) S believes that broccoli causes flat feet

is to be understood as having the structure,

(2) S believes some •Broccoli causes flat feet•,

the dot quotes serving both to form what is grammatically a common noun and to ostend or demonstrate the linguistic token that they enclose. (Strictly speaking, the ostention is deferred, as we shall see.) Alternatively, perhaps a clearer representation of (1) would be

(3) S believes one of those.



Belief, then, is a two-place relation holding between a person and an inscriptional token that falls under a certain type; a person has a belief when he or she bears that relation to such a token.

This grammatical/semantic account of belief ascription also makes a nice ontological/psychological theory of believing (or so I have argued in (1981c)), one that fits nicely into a functionalist theory of mind of the sort I favor. For there is good reason to think that people sometimes host inner episodes, which I construe as brain episodes, that have the sorts of structure that we associate with bits of public language (Sellars (1956); Fodor (1975), (1978); Field (1978)). E.g., a single such episode might have subject-predicate form, i.e., might consist of the tokening of a mental representation of an object X concatenated with the tokening of a representation that expresses the property of being F; in so doing the episode would count as a •X is F•, just as would a public utterance.

But in virtue of what would a brain episode or its parts be a representation of X or of F-ness or of anything else? How is it possible for purely physical events to have "aboutness" or intentionality? That depends on what we think it takes for anything to have aboutness, for anything to refer or be intentional, and on whether our answer to that question is extrapolable from paradigm cases in public language, say to the more recondite case of brain activity. This is an old, old story, first told in modern times by Sellars and subsequently amplified and sophisticated by Jerry Fodor (1975), Hartry Field (1978), and others (see also Harman (1973) and Rosenberg (1974)). A lesson of the past ten years' work in the theory of reference is that what mediates linguistic aboutness is causality--causal chains of certain roughly specifiable types,[13] and/or reliable functional links consisting

in linguistic expressions' specific sensitivity to their referents due to conditioning (cf. Dennett (1969), Harman (1973), Bennett (1976), Stalnaker (1979b), Wimsatt (1968), Fodor (1982), and others). However the details of these connections might go, the important points to note are (i) that the connections are _naturalistic_ (they are real relations to be found in nature), and (ii) that other physical items besides public linguistic tokens can bear them to objects--brain events, in particular, can have the appropriate sorts of etiologies and can be conditioned responses showing specific sensitivity to particular objects and types of object. Thus, we have every reason to think, or at least to hope, that brain events can be intentional in (almost) just the way that bits of language are.

So: To judge or believe that P[14] is to bear the "belief" relation to an inner representation whose syntactic/semantic structure is analogous to that of the sentence that replaces "P" (more on which analogousness shortly). The "belief" relation itself is a distinctive _functional_ relation, consisting in the representation's playing a certain type of functional role, i.e., doing a certain type of administrative job within the functional hierarchy that is the believer himself. Obviously the role that is characteristic of beliefs as opposed to desires, intentions, and so on has to do with storage, with mapping, and otherwise with serving as a guide to action.[15]

There is a great tendency in the recent literature[16] to overstate the commitments of this representationalist theory, and indeed to caricature it unmercifully, as if the representationalist were suggesting that inside each believer's head is a tiny blackboard with all the believer's stored representations written in chalk, or that evil, politically motivated scientists might be able to spy inside our heads with their cerebroscopes and report our innermost convictions to the Thought Police. I have disavowed these straw-man interpretations in earlier works (see particularly (1981c), footnotes 4 and 6), and tried to show why no such absurd consequences in fact flow from the representationalist view, but let me say another word about the matter here, borrowing an example from Stephen Stich.[17] Many people balk at representationalism because they hold that _higher animals_ and pre-linguistic children have beliefs. If dogs, say, have beliefs but no language, then how can belief be a matter of bearing a functional relation or any other relation to a quasi-linguistic item? Isn't that idea just silly in the case of dogs?

One possible answer to this would be to suppose that the relation is a counterfactual one. That answer has possibilities, but I think there is a much more straightforward one: Suppose Lassie is digging at a particular spot because she believes it is where she buried a bone earlier in the day. Are we really to suppose that no inner state of Lassie keys on

the spot in question and guides her to it (and would have guided her to it even if various obstacles had been placed in her path)? Are we to suppose that no inner state of Lassie is specifically a _bone_ response and is causally connected in this instance to the state that keys on the spot? That seems to me very unlikely. And all the representationalist is saying is that some such states of Lassie's central nervous system are performing those functions and related ones; that is all it takes for Lassie to accept a ·Bone here·.[18]

A further reason why people balk at representationalism, is that they think the representationalist is making the obviously false claim that the intentionality of public linguistic items is logically or ontologicaaly prior to that of beliefs and thoughts. This impression is encouraged by talk (engaged in even by representationalists) of "sentences in the head," "brain writing," "inscriptionalism," and the like. As I have pointed out before ((1981c), footnote 4), this talk is misleading in exactly the way we would mislead if we were to speak of public linguistic tokens as being "thoughts out in the open air." Thoughts are representors in the brain, made of brain stuff; bits of language are representors out in the open air, made of sound waves or ink marks; tape recordings are representors made of patterns of magnetism; and so on. We seek a general theory of representation that covers all types of representor.

So far none of what I have said about intentionality is at all original. But now it is time to unveil the gimmick I shall use in attacking the paradox of naming. Let us turn back to the mechanics of dot quotation.

5.

Dot quotes, and plural demonstratives of the sort that occur in (3) above, _classify_ the tokens they ostend. Now, how do they do that? How do we determine the extension of '... is a ·Broccoli causes flat feet·,' or tell when some token of a very different shape _is_ "one of _those_"? Philosophers have offered varying advice about this. Sellars himself counts an item as a ·Broccoli causes flat feet· just in case it plays approximately the same _inferential_ _role_ (augmented by "language-entry" and "language-exit" norms) within its own surrounding conceptual framework as the English sentence "Broccoli causes flat feet" plays within English. Davidson merely introduces a relation of "samesaying" without further comment, but given the semantical format for which he is famous, we might suppose he would count something as a ·Broccoli causes flat feet· just in case it has the same truth condition as does the corresponding English sentence; or we might insist on a stronger intensional isomorphism, perhaps based on the procedure by which the truth-condition is computed. Other

individuative methods are possible also, and the multiplicity of possibilities is what is going to help us resolve our paradox.

Let us ask ourselves, as Kripke does by implication in (1979): just what does poor S believe about Cicero and about Tully. Does he believe that Cicero was bald? Less tractably, does he believe that Tully was bald? There seems no obstacle to answering "Yes" to the first of these questions, but intuitions divide sharply over the second. Impressed with the felt rigidity of names, I have in previous writings[19] inclined toward maintaining that S does believe that Tully (that very person, whom S himself happens to be calling 'Cicero') was bald, even though S himself would never express his belief in that way, and that S does hold contradictory beliefs, though he has no way of detecting this. (He also believes that Cicero (=Tully) was not bald; another contradiction.) Other philosophers are more impressed by the commonsensical argument that if S sincerely, coherently, and lucidly asserts "Tully was not bald," then S believes that Tully was not bald (and doesn't believe that Tully is).

I now think an irenic solution is available. Both intuitions should be respected, and thanks to the Sellarsian semantics they can be, in a way that has independent motivation as well. Recall that we have a choice of classificatory schemes to impose on our proffered exemplars, i.e., of methods for telling which tokens count as ·[so-and-so's]·s and which don't. Proponents of the Sellarsian approach have always tended to suppose that some one such scheme is the correct scheme. But suppose we reject that presumption instead. There is considerable independent evidence[20] that at least two disparate individuative schemes are used on different sorts of ascription-occasions, depending on the context and the conversational point of the ascription. Sometimes we follow Sellars himself in classifying our tokens according to their inferential roles, or perhaps I should say, according to their *computational* roles *à la* Jerry Fodor,[21] when what interests us are the tokens' causal properties and the explanation and prediction of behavior. (The lesson of "methodological solipsism" as I understand it is that a belief's computational surface and (emphatically) not its truth-condition or other semantical property is what determines its causal contribution.[22]) But sometimes we do individuate beliefs semantically, e.g., according to sameness of truth-condition, when we are more interested in truth-related aspects of belief, such as the believer's success in achieving his goals, or the success of our use of his belief reports as authoritative evidence for our own beliefs.

Our two-scheme hypothesis affords us a properly equivocal resolution of our dilemma and a likewise equivocal answer to Kripke's type of hard question. First, does S irrationally

believe a contradiction? When he makes his celebrated utterance, he does so partly as a result of having affirmed a mental version of the sentence he utters; he has accepted a •Cicero was bald and Tully was not•. The Millian, or even the proponent of names' rigidity, will insist that this amounts to accepting a •Cicero was bald and Cicero was not•, since if names are Millian or at least rigid, the names 'Cicero' and 'Tully' make exactly the same truth-conditional contribution, and the two tokens just displayed between dot quotes express exactly the same belief in the sense of being true at just the same worlds (and for the same reason). But is *S*'s mental analogue of "Cicero was bald and Tully was not" really a •Cicero was bald and Cicero was not•? That depends on our choice of individuative schemes for dot-quoted items. If we appeal to the semantical scheme, we find it does count *S*'s mental token as a •Cicero was bald and Cicero was not•, and in that sense *S* does believe a contradiction. But *option* (*I-b-ii*) *is in force--in the present sense*, there is nothing irrational about believing a contradiction. For *S* has no syntactic way of detecting his semantical anomaly; he cannot deduce from his mental token anything he could recognize as a contradiction. The relevant contents of his head are analogous to an uninterpreted formal calculus equipped with rules of natural deduction. In the absence of "*a* = *b*" as an axiom, it would be positively *ir*rational to infer by substitution of '*b*' for '*a*', even if on some preferred interpretation that *could* be supplied by an external observer, '*a*' and '*b*' are assigned the same referent.

Now, what about appeal to the computational individuative scheme? Computationally speaking, *S*'s mental token of "Cicero was bald and Tully was not" does *not* count as a •Cicero was bald and Cicero was not•, since the representations associated respectively with the names 'Cicero' and 'Tully' play obviously distinct inferential and computational roles for *S*, and accordingly distinct behavior-causing roles.[23] From "Cicero was bald and Tully was not," e.g., *S* would infer "Cicero was bald," "Tully had a property that Cicero didn't," "Cicero and Tully were two different people," etc., and would *not* infer "Tully was bald" or "Tully existed." But from "Cicero was bald and Cicero was not," *S* would either start inferring every sentence he could think of, or go into cognitive spasm of some sort (given a generous helping of downward causation, *S*'s circuitry might turn black and give off smoke). Thus, in this sense (according to this individuative scheme), *S* does *not* believe that Cicero was bald and Cicero was not, even though he believes that Cicero was bald and Tully was not. It is probably psychologically impossible to believe an explicit contradiction *computationallly individuated*.

Several authors have, I think, glanced off the distinction between our two individuative schemes, without entirely

realizing that that was what they were doing. Hartry Field, Jerry Fodor, David Lewis, Stephen Stich, and most recently John Perry have hinted at it.[24] But, interestingly, they have in effect taken sides (different sides) on which of the two schemes is correct, or at least on which is vital to the concept of belief and which negligible. Fodor and Lewis assume that beliefs are essentially causal entities invoked to explain behavior and that their semantical properties are by the way, while Stich and Perry insist that the truth-values of beliefs (and the reliability of informants) are what matter and that explaining behavior isn't so important after all. Now, this seems to me a funny sort of thing to _quarrel_ about. Sometimes we're interested in explaining and predicting behavior; at other times we're interested in truth and reliability. Which of these interests is _objectively paramount_ seems to me an idle question. And if my Sellarsian semantics is right, our language affords us a pragmatic choice in belief ascription, that matches our pair of alternative interests nicely.

Does _S_ believe that Tully was bald? On our two-scheme hypothesis the answer is, quite properly, "_Yes and no_." On the computational scheme, _S_'s mental analogue of "Cicero was bald" does not count as a ·Tully was bald·; _S_'s mental "Cicero" and his mental "Tully" play entirely different computational roles, even though they are in fact grounded in (are representations _of_) one and the same person. On the semantical scheme, _S_ does believe that Tully was bald, since that scheme does count his inner analogue of "Cicero was bald" as a ·Tully was bald·. Thus, by providing for an ambiguity in the reference of the plural demonstrative underlying the complementizer 'that', our Sellarsian account is able to predict and explain our preanalytic uncertainity and disagreement about "_S_ believes that Tully was bald."[25]

6.

Let's briefly revisit the topic of _abbreviation_, for it too is illuminated by our two-scheme hypothesis. In section 1 we rejected as preposterous the suggestion that an abbreviative procedure literally occurs whereby Russellized descriptions are squished into names. We have also wanted to reject (II-a), the Russellian option, _tout court_, and to deny that names are even _equivalent_ to flaccid descriptions in any semantical sense. Yet something funds the Russellian intuition, and that something is brought out by what we might call the "spot-check test." (I use this in teaching Russell to undergraduates.) Suppose we overhear Jones using a certain name, say, "Wilfrid Sellars," and we want to test his grasp of his own ideolect. Whipping around, we demand, "Who do you mean by 'Wilfrid Sellars'? Quick, now!" Subjected to this test, Jones will doubtless produce a description, as in "Oh, I

mean the famous philosopher at Pittsburgh who wrote 'Empiricism and the Philosophy of Mind'." This phenomenon suggests that the ability to produce a description on demand is constitutive of competence in the use of a name, and Russellians might take it as proving the synonymy of names with descriptions. (That it does not prove that is what Kripke showed in enforcing his now well-known distinction between a description's actually "fixing the sense" of a name and the description's merely "fixing [the name's] reference.") But now we are in a position to explain the results of the spot-check test without incurring the consequence that (II-a) is correct.

A Mentalese name, as used by a person at a time, can have the same computational role as a description for the person at that time. Thus it can be "equivalent to" the description *modulo* the computational individuative scheme, and two of the subject's beliefs, alike except that one involves the name while the other involves the description, will be computationally and hence causally similar for the subject at the time--the two beliefs will be functionally equivalent for all practical purposes without being semantically equivalent at all. Why should names ever share the functional roles of descriptions in this way? I think the answer must be that functionally speaking, names are something like labels on files, or perhaps more like tabs on files, where each file is a store of contingent information associated with the name. A tokening of the inner name calls up the most salient information in the file (perhaps tokening just is the calling up of that information), and that's why we feel that particular uses of names are "backed by" particular bodies of descriptive material. It is also why the spot-check test works. (It is also, I should think, why identity statements involving names are "informative" despite the triviality of their semantical contents: To accept an identity statement is to merge files.[26] Or at least, when an identity statement is accepted files get merged; and cognitive capacities thereby usefully consolidated.)

Names to not abbreviate descriptions in any semantical sense at all. They just share computational roles with descriptions from time to time.

7.

I have argued in (1981c) that the two-scheme hypothesis also solves the problem of self-regarding attitudes, and that it succeeds in sorting out some puzzles and misconceptions about "methodological solipsism." I won't repeat those discussions here, but I shall close by trying to rebut an objection connected with the second.

If it is our computational individuative scheme that pertains directly to the explanation of behavior, then it is the scheme that would officially be mobilized by cognitive

psychologists, at least insofar as their official concern is the explanation of behavior. Thus the psychologists will forswear the semantical individuative scheme, and will not characterize beliefs according to their semantical properties. (A belief's propositional content--identifiable, let's say heuristically, with the set of worlds at which the belief is true--is irrelevant to the belief's psychological role; surprising, but obvious enough when you think about it.) Now, Stephen Stich has taken these facts to prove that the psychologists will no longer be talking about beliefs at all and in that sense are no longer cognitive psychologists. I have tried to show that in view of the availability and the probity of the computational individuative scheme, this doesn't follow; but it does leave us with the problem of saying how the psychologists are to talk. I suggested in (1981c) (footnote 11) that cognitive scientists would have to describe beliefs using "some appropriately neutral syntactic code," by which I intended, adhering to our usual computer analogy, some kind of machine language. This is a perfectly tenable theoretical suggestion, so far as I can see, but doesn't help much considered as practical advice. For not even computer scientists, or computer operators, adhere to machine-code description in dealing with their machines from day to day; and we have no "machine language" for the brain in any case.

A better suggestion would be simply to fix the parameters that effect the valuations for the subjects' inner representations. That is, allowing that the semantic values of the representations' constituents are determined in part by causal, historical, and social antecedents outside their owners' heads, there still exists a valuation function V from contexts and representations to propositional objects that both determines the representations' propositional objects given the relevant sets of intra- and extracalvarian factors, and serves as our interpretive manual for the subjects' speech and behavior. V is complicated and subtle, but plainly we have at least an effective working grasp of V (however incomplete) as part of our linguistic competence; we are usually able to tell what other people are talking and thinking about. Now, given that our psychologists grasp V and are also fairly well in command of the contextual factors that are among V's arguments, they can simply describe their subjects' inner states modulo that set of factors, which set is part of their shared background information in any ordinary research context. (In just the same way, a computer operator describes his machine as just having "computed Robin Roberts' batting average" or whatever, not because reference to Robin Roberts occurs completely inside the machine's skin, which it obviously doesn't, but because he is filling in contextual information about the etiology and social background of the relevant computational states of the machine.) The only caveat is that the

psychologists must not forget that that is what they are doing, and lapse into supposing that reference to Robin Roberts or to anything else is somehow determined by what is inside their subjects' heads. There is no harm so as long as they and we remember that genuineness of reference makes no difference to the homunculi in the driver's seat.

All this presupposes that the psychologists have epistemic access of some sort, however indirect, to the subjects' representàtions, and that is highly problematic. But it was problematic anyway. I see no special difficulty of the sort suggested by Stich's argument.[27]

NOTES

1. E.g., see Schiffer (1978).

2. Kripke (1972). See also Kaplan (1979) and elsewhere, and Donnellan (1976).

3. There are a number of other arguments that I find unconvincing. I discuss them in (1980a); see also McKinsey (1978).

4. Kripke (1979). Other recent criticisms of (I) may be found in Schiffer (1978), in Ackerman (1979b), and particularly Baker (in press).

5. I have tried to show this in Boër and Lycan (1975), sec. III, in Boër and Lycan (1980), and in Lycan (1978). See also Marcus (1981).

6. There is a further difficulty of how to fix the reference of '@' itself, pointed out by David Austin in (1979).

7. This formulation is very closely related to Kripke's "Puzzle about Belief" (1979), though the case he concentrates on is more complicated. The solution I go on to propose would work for his case also.

8. E.g., see Frege's footnote to p. 58 of (1966).

9. I think this is generally agreed; however, Michael Dummett contests it in (1973).

10. Here is another way of putting the same point: Hintikka's semantical treatment of (A) requires that 'Cicero' name one world-line and 'Tully' name a diverging one. Therefore there is a nonactual world w at which Cicero

and Tully are distinct. It follows from this by the definition of 'rigid' that either 'Cicero' or 'Tully' is nonrigid in (A), and given our assumption that the two names have the same (type of) semantical status, they are both flaccid. In addition, corresponding to each of the two world-lines there is a function from worlds to sets: at each world the function spits out the unit set of the relevant individual manifestation. A function from worlds to sets determines a property, however complex or esoteric the property might be. Since our 'Cicero' and 'Tully' world-lines diverge, they determine distinct properties, at least one of which is lacked by our common referent at some world, which is to say that they determine different contingent properties of the referent. Thus, each name turns out to be semantically equivalent to the description whose matrix expresses the relevant contingent property.

(The foregoing is one natural reading of Hintikka's position. I should emphasize that his view could easily be modified in such a way as to avoid this consequence, though I think at the cost of collapsing it into Plantinga's option. Hintikka's world-line apparatus itself raises deeper skeptical questions about our näive "rigid"/"flaccid" distinction, also; cf. Kraut (in press).)

11. Two other recent theorists whose purported solutions to our dilemma collapse (I think) into already existing options are Diana Ackerman ((1979a), (1979b)) and Robert Stalnaker ((1978), (1979a)). Ackerman argues by way of a very demanding substitutivity test that although names undoubtedly have connotations, the connotations they do have are not shared by any descriptions. However, by her standards, many or most natural-language expressions have "connotations" that are not shared by any other expressions, and our interest shifts to her more permissive notion of the correct "analysis" of names; it turns out that for her names are equivalent to world-indexed descriptions that incorporate a causal-chain theory of reference. A name N, she says, is equivalent to (roughly) ⌜the entity bearing R in @ to this very use of 'N'⌝, where 'R' denotes the appropriate sort of causal-historical relation. Thus, if I understand her correctly, she joins Plantinga in attributing a contradictory belief to our friend S.

Stalnaker proposes an ingenious appeal to two-dimensional modal logic, in an effort to determine what proposition is in fact believed by someone in S's position; Stalnaker is reluctant on the grounds of charity to charge S with self-contradiction. His independent moti-

vation for his charitable interpretation of S's utterance is highly original (I discuss it in (1980b)), but the interpretation itself is not--so far as I can see, he portrays S as ending up believing just the propostion that Russell would have said S believed. (However, Stalnaker's treatment applies only within intensional operators; he agrees with Kripke that names occurring outside such contexts are rigid.)

12. I failed to appreciate this fully in sec. III of Boër and Lycan (1980), and spent only one page (451) on inadequate psychological handwaving.

13. The best developed causal-chain theory I know is that of Devitt (1981); see also Boër (in press).

14. For purposes of this paper I shall restrict discussion to occurrent or episodic belief. I am not happy with the account, or non-account, of "tacit" belief that I relied on in (1981c), and am in the process of replacing it.

15. On the teleological nature of functional characterization, see my (1981a), (1981b) and (1981c).

16. Perhaps the best example of this is Dennett (1978).

17. Stich (1979). I emphasize that I am borrowing only the example and not the ingenious argument that goes with it; that argument deserves considerable discussion in its own right.

18. Robert Stalnaker begins his (1979b) by contrasting the Sellarsian inner-speech paradigm with an opposing picture that shows the intentionality of thought as being prior to that of speech in any form, and he announces his intention to defend a version of the latter at the expense of the former. Ironically, though, the account that emerges in his concluding section is as nice a specimen as one could want to the naturalistic sort of representationalism I am advocating here.

    Incidentally, we must concede to Stich that there can be no straightforward identification of Lassie's inner predicate with our concept "bone." My present use of 'bone' in describing Lassie's belief is less than fully accurate (but it's fast).

19. See the works cited in footnote 5.

20. I have in mind here Hector-Neri Castañeda's data ((1966), (1967)) involving self-regarding attitudes, and

the Twin-Earth cases adduced by Hillary Putnam (1975) and by Jerry Fodor (1980) in support of "methodological solipsism." I have discussed these in (1981c).

21. These schemes are similar but distinct. For Sellars, linguistic roles (as marked by rules of assertibility, rules of inference, and the like) are essentially normative, while computational roles for Fodor are (I take it) more purely causal. My own preference is for a cautious fusion of the normative and the causal in the form of the teleological; cf. footnote 15.

22. See Fodor (1980), Stich (1978), and Lycan (1981c).

23. What are "the representations associated respectively with" particular public names such as 'Cicero' and 'Tully'? Whatever inner states, I suppose, are typically activated by <u>S</u>'s hearing those names tokened and/or which typically figure in <u>S</u>'s production of those names.
    The only explicitly Sellarsian treatment of proper names that I know of, incidentally, is Jay Rosenberg's excellent (1978); he does not explore the two-scheme hypothesis.

24. Field (1978), pp. 48, 51; Fodor (1980), p. 67; Lewis (1981), pp. 288-289; Stich (1978); Perry (1982). Since writing the main body of this paper I have also seen Brian Loar's very interesting book (1981), on p. 117 of which he makes almost exactly the point I am pushing here.

25. Incidentally, the two-scheme approach preserves Kripke's "disquotation" principle (1979) to the effect that <u>S</u> will sincerely and competently assert "P" only if <u>S</u> believes <u>that</u> P. For what that's worth; I have some independent doubts about the principle.

26. A similar view of identity statements is put forward in Lockwood (1971).

27. This paper began life as a 1979 memo to the Herbert Heidelberger Task Force on the Propositional Attitudes, at the University of Massachusetts. I am grateful to Heidelberger, to his co-founder Murray Kiteley, and to Lynne Rudder Baker and David Austin for many lengthy discussions and comments; I also thank Robert Kraut and Steven Boër.

REFERENCES

Ackerman, Diana: 1979a, 'Proper Names, Essences and Intuitive Beliefs', Theory and Decision, 11.

Ackerman, Diana: 1979b, 'Proper Names, Propositional Attitudes and Non-Descriptive Connotations', Philosophical Studies, 35.

Austin, David: in press, 'Plantinga's Theory of Proper Names', Notre Dame Journal of Formal Logic, forthcoming.

Baker, L.R.: in press, 'Underprivileged Access', Nous, forthcoming.

Bennett, Jonathan: 1976, Linguistic Behavior, Cambridge University Press, Cambridge.

Boër, Steven: 1984, 'Substance and Kind: Reflections on the New Theory of Reference', Analytical Philosophy in Comparative Perspective, D. Reidel, Dordrecht.

Boër, Steven and W.G. Lycan: 1975, 'Knowing Who', Philosophical Studies, 28.

Boër, Steven and W.G. Lycan: 1980, 'Who, Me?', Philosophical Review, 89.

Castañeda, H.-N.: 1966, '"He": A Study in the Logic of Self-Consciousness', Ratio, 8.

Castañeda, H.-N.: 1967, 'Indicators and Quasi-Indicators', American Philosophical Quarterly, 4.

Davidson, Donald: 1968, 'On Saying That', Synthese, 19.

Dennett, D.C.: 1969, Content and Consciousness, Routledge and Kegan Paul, London.

Dennett, D.C.: 1978, 'Brain Writing and Mind Reading', in Brainstorms, Bradford Books, Montgomery, Vt.

Devitt, Michael: 1981, Designation, Columbia University Press, New York.

Donnellan, Keith: 1976, 'Speaking of Nothing', Philosophical Review, 85.

Dummett, Michael: 1973, Frege: Philosophy of Language, Harper and Row, New York.

Field, Hartry: 1978, 'Mental Representation', _Erkenntnis_, 13.

Fodor, J.A.: 1975, _The Language of Thought_, Crowell, New York.

Fodor, J.A.: 1978, 'Propositional Attitudes', _Monist_, 61.

Fodor, J.A.: 1980, 'Methodological Solipsism Considered as a Research Strategy in Cognitive Psychology', _Behavioral and Brain Sciences_, 3.

Fodor, J.A.: 1982, 'Psychosemantics, or: Where Do Truth Conditions Come From?', unpublished xerox.

Frege, Gottlob: 1966, 'On Sense and Reference', in Geach and Black (eds.), _Translations from the Philosophical Writings of Gottlob Frege_, Blackwell, Oxford.

Harman, Gilbert: 1973, _Thought_, Princeton University Press, Princeton.

Hill, C.S.: 1976, 'Toward a Theory of Meaning for Belief Sentences', _Philosophical Studies_, 30.

Hintikka, K.J.J.: 1969, 'Semantics for Propositional Attitudes', in _Models for Modalities_, D. Reidel, Dordrecht.

Hintikka, K.J.J.: 1975, _The Intentions of Intentionality and Other New Models for Modalities_, D. Reidel, Dordrecht.

Kaplan, David: 1979, 'Dthat,' in French, Uehling and Wettstein (eds.), _Contemporary Perspectives in the Philosophy of Language_, University of Minnesota Press, Minneapolis.

Kraut, R.: in press, 'Hintikka's Ontology', in R.J. Bogdan (ed.), _Profiles: Jaakko Hintikka_, D. Reidel, Dordrecht.

Kripke, Saul: 1972, 'Naming and Necessity', in Davidson and Harman (eds.), _Semantics of Natural Language_, D. Reidel, Dordrecht.

Kripke, Saul: 1979, 'A Puzzle About Belief', in A. Margalit (ed.), _Meaning and Use_, D. Reidel, Dordrecht.

Lewis, David: 1981, 'What Puzzling Pierre Does Not Believe', _Australasian Journal of Philosophy_ 59.

Loar, Brian: 1981, _Mind and Meaning_, Cambridge University Press, Cambridge.

Lockwood, Michael: 1971, 'Identity and Reference', in Munitz

(ed.), Identity and Individuation, New York University Press, New York.

Lycan, W.G.: 1978, 'Referential Opacity Explained Away', talk delivered at the University of Sydney.

Lycan, W.G.: 1980a, 'Kripke's Arguments Against the View that Proper Names Abbreviate Descriptions', unpublished ditto.

Lycan, W.G.: 1980b, 'Thoughts on Stalnaker's Semantics for Belief', unpublished ditto.

Lycan, W.G.: 1981a, 'Form, Function, and Feel', Journal of Philosophy, 78.

Lycan, W.G.: 1981b, 'Psychological Laws', Philosophical Topics, 12.

Lycan, W.G.: 1981c, 'Toward a Homuncular Theory of Believing', Cognition and Brain Theory, 4.

Marcus, Ruth B.: 1981, 'A Proposed Solution to a Puzzle About Belief', Midwest Studies in Philosophy, 6.

McKinsey, Michael: 1978, 'Kripke's Objections to Decripition Theories of Names', Canadian Journal of Philosophy, 8.

Perry, John: 1982, 'Belief and Opacity in Situation Semantics', talk delivered to the University of Massachusetts' Workshop on Propositions, Propositional Attitudes, and Finite Representability.

Plantinga, Alvin: 1978, 'The Boethian Compromise', American Philosophical Quarterly 15.

Putnam, Hilary: 1975, 'The Meaning of "Meaning"', Minnesota Studies in the Philosophy of Science vol. 7, University of Minnesota Press, Minneapolis.

Rosenberg, Jay: 1978, 'Linguistic Roles and Proper Names', in J.C. Pitt (ed.), The Philosophy of Wilfrid Sellars: Queries and Examinations, D. Reidel, Dordrecht.

Rosenberg, Jay: 1974, Linguistic Representation, D. Reidel, Dordrecht.

Schiffer, S.: 1978, 'The Basis of Reference', Erkenntnis 13.

Sellars, Wilfrid: 1956, 'Empiricism and the Philosophy of Mind', Minnesota Studies in the Philosophy of Science vol.

1, University of Minnesota Press, Minneapolis.

Sellars, Wilfrid: 1963, 'Some Reflections on Language Games', in Science, Perception, and Reality, Routledge and Kegan Paul, London.

Sellars, Wilfrid: 1967, Science and Metaphysics, Routledge and Kegan Paul, London.

Sellars, Wilfrid: 1969, 'Language as Thought and as Communication', Philosophy and Phenomenological Research 29.

Sellars, Wilfrid: 1973, 'Reply to Quine', Synthese 26.

Stalnaker, R.: 1978, 'Assertion', in P. Cole (ed.), Syntax and Semantics vol. 9, Academic Press, New York.

Stalnaker, R.: 1979a, 'Semantics for Belief', unpublished xerox.

Stalnaker, R.: 1979b, 'Thoughts', unpublished xerox.

Stich, Stephen P.: 1978, 'Autonomous Psychology and the Belief-Desire Thesis', Monist 61.

Stich, Stephen P.: 1979, 'Do Animals Have Beliefs?' Australasian Journal of Philosophy 57.

Wimsatt, W.C.: 1968, 'Purposiveness and Intentionality in Nature', unpublished xerox.

Steven E. Boër

SUBSTANCE AND KIND: REFLECTIONS ON THE NEW THEORY OF REFERENCE

1. INTRODUCTION

The traditional doctrine of intension and extension, enshrined in nearly every elementary logic text (under the heading "Definitions"), is a litany of comfortable words. The extension of a (singular or general) term is, we are told, the class of things "it applies to", and its intension is that associated feature of the term in virtue of which it so applies--roughly, an assemblage of non-question-begging necessary conditions which are jointly sufficient for the term's application. For most purposes, it is said, talk about the "meaning" of a term may be construed as talk about its intension: in particular, "knowing the meaning of a term" may be construed as knowledge of its intension, i.e., having a "mental representation" of the necessary-and-sufficient condition for its application. This beguilingly simple picture has undergone much technical refinement in the hands of logicians, whose formal implementations have contributed to its long and unbroken reign among semantic theorists.

Yet there is a disturbing intuition, voiced of late by Kripke, Putnam, and others, that the Emperor has no clothes, that the simple picture alluded to above cannot be the *whole* truth about the referential mechanisms operative in natural languages (though it may well be, and probably is, a correct account of how *some* linguistic items function). Here am I, the owner of five cats and a frequent user of 'cat'; surely I would count myself a competent user of this word. After all, I have (so I believe) no practical trouble when it comes to *identifying* cats, and I have a fair amount of empirical knowledge about their characteristics. But when I search for my mental representation of the intension of 'cat', I come up empty-handed. The best I can do is to describe the appearance and behavior of typical cats and to point to some alleged exemplars. But none of this adds up to a set of individually necessary and jointly sufficient conditions for something's being a cat: there may be atypical cats, who look and behave quite differently; and things which look and behave as described may not be cats. For all I know, there may be cat-like marsupials which pass for cats in something like the way in which koalas pass for bears; and even my own five might be

B. K. Matilal and J. L. Shaw (eds.), Analytical Philosophy in Comparative 1

such marsupial masqueraders! If what determines the extension of 'cat' in my mouth is my mental record of the intension of 'cat', then, in default of the latter, it seems that I must conclude that in all these years of using 'cat' I have not (save perhaps in purely ostensive situations) managed to talk about anything at all, let alone cats. Indeed, since virtually every English-speaker finds himself in a similar circumstance anent 'cat'--not to mention 'dog', 'horse', 'human being', etc.--it threatens to become a mystery how people have ever managed to talk about the world at all! At this point there is surely some ground for suspecting a flaw in the traditional doctrine. It is against the background of intuitive considerations like these (and related considerations for the case of proper names) that Kripke, Putnam, and others have developed the so-called New Theory of Reference, which attempts to remedy what it diagnoses as the fundamental mistake in the orthodox view.

In this paper I shall be concerned with the application of the New Theory of Reference to the use of mass nouns and count nouns as substance and kind words, as carried out in Kripke [1972], Putnam [1970b], [1973], [1975a], [1975b], and as extended in Goosens [1977] and Teller [1977]. After sketching the development of a "minimal" form of the New Theory for such uses of mass and count nouns, I shall address a number of criticisms which have accumulated in the literature (and which have for the most part gone unanswered). By progressive refinements of the minimal theory in the face of these criticisms, I shall attempt to provide a version of the New Theory of Reference for mass and count nouns which both conforms to the spirit of minimal theory and is proof against extant objections.

2. DEVELOPMENT OF THE MINIMAL THEORY

The intuitive considerations adduced above were, of course, only meant to be suggestive, not probative; there are plenty of rejoinders open to devotees of the traditional doctrine (let's call them 'Intensionalists'). But there is one point on which even the most rabid Intensionalist must agree with the proponents of the New Theory: whether or not general terms like 'cat' have associated intensions, a person can be a competent user of the term (can succeed in using it to talk about cats of whatever) on a given occasion even though <u>he</u> fails on that occasion (or perhaps even on any occasion) to possess any definition of the term via some non-question-begging necessary-and-sufficient condition. To concede this point is not really to concede very much, since it does not follow that no one ever has to possess such a definition. Where singular terms are concerned, the Intensionalist usually makes this concession without prodding, allowing that one

person may "borrow" another's singular reference by intending to refer to whatever that person was referring to with the term in question. But the buck cannot be passed indefinitely and must sooner or later stop with a possessor of an appropriate intension. By parity of reasoning, then, it may be allowed that I am able to use 'cat' to speak of cats because I have "borrowed" this ability from others who have it in virtue of grasping the posited intension of 'cat'. If we think of these lucky savants as "the experts", we can readily see that Putnam's "Hypothesis of the Division of Linguistic Labor" is really quite neutral as between Intensionalism and the New Theory of Reference. For the Hypothesis amounts only to the claim that

> Every linguistic community ... possesses at least some terms whose associated 'criteria' are known only to a subset of the speakers who acquire the terms, and whose use by other speakers depends upon a structured cooperation between them and the speakers in the relevant subsets. (Putnam [1975a], p. 228)

(Page references to Putnam [1962] - [1975b] are to the reprinted versions in Putnam [1975c] and [1975d].)

If the Division of Linguistic Labor were the only point at issue, there would be nothing new in the New Theory of Reference. But there is more to our initial story than that. For the "experts" in question typically come <u>after</u> the fact, not beforehand. The word 'cat' (or its linguistic ancestor) was in common usage long before there were any acknowledged experts on cathood such as zoologists or veterinarians. People talked about cats long before there were any "authorities" on the criteria of cathood from whom to borrow (presumably this is not true for theoretical terms, about which we shall have something to say in due course). At this juncture, the Intensionalist begins to feel the heat. Whatever the <u>current</u> intension of 'cat' is, it is presumably fairly complicated--a matter, say, of genetic and/or evolutionary properties--and known only to the scientifically educated few, from whom we laymen are entitled to borrow. <u>Pre</u>scientific uses of 'cat' obviously could not have had <u>this</u> intension, yet they must have had, or been linked by borrowing to uses which had, some intension(s) or other. From this it is but a short step to the view that 'cat' has continually <u>changed</u> <u>its</u> <u>meaning</u> (intension) over time--and very probably its extension as well--beginning with some "primitive" intension to secure initial reference and gradually acquiring new intensions in light of new information and experiences.

The awkwardness of the Intensionalist's picture is that it flies in the face of common sense, which, for better or

worse, has it that people have been using 'cat' (and its ancestors) for centuries to talk about one and the same sort of thing--cats--and that what has changed over time is merely the content of our _beliefs_ _about_ cats. On the Intensionalist's picture, however, such continuity of reference would be nothing short of a miracle. The best the Intensionalist dare hope for is an historical series of intensions determining a corresponding series of distinct but pairwise overlapping extensions, for he must regard each shift in beliefs formulated with 'cat' as (at least potentially) a shift in its intension.[1] At this juncture the New Theory of Reference reminds us of a simple but important fact: we often answer questions of the form 'What is an _F_?' by pointing to (what we take to be) some _F_-things and saying '_These_, _and_ _things_ _of_ _the_ _same_ _kind_, are _F_s' or '_That_ _sort_ of thing is an _F_'. Since cats are nicely observable, this fact about questions and answers suggests that an extension for 'cat' might be fixed, without the need for any analytical definition, via some ceremony employing a formula like 'Let 'cat' apply exactly to those things which are of the same kind as _these_ things', the referent(s) of the demonstrative being provided through ostension of "paradigm" cats.

Two problems of interpretation immediately arise, one regarding the indexical component in the envisaged reference-fixing formula and the other concerning the ingredient phrase 'of the same kind as'. First of all, given a paradigm cat, X, who or what determines when something is of the same kind as X? Earlier we spoke of "experts"--why not let them decide? The answer is obvious: if experts are just locally acknowledged authorities, there is nothing to prevent (a) the experts at a given time being _wrong_ in their classifications, (b) experts at a given time disagreeing among themselves, or (c) experts at different times delivering wildly different verdicts. Any of (a) - (c) would be destructive of the commonsensical unity and continuity of reference so dear to the heart of the New Theorist. Clearly somebody must have the last word, if not the expert of today then perhaps future experts. In line with this idea, it is often proposed that the burden falls upon "Final Science", construed either realistically or in Peircean terms. Let us, then, swallow the notion of Final Science for expository purposes and take 'is of the same kind as' to be tacitly qualified by the rider, 'by the lights of Final Science'.

Zemach [1976] points to our second exegetical problem. If the extension of 'cat' is thought of as an equivalence class collectively generated by the speech community's paradigm cats under the aforementioned sameness-of-kind relation, we face the embarrassing possibility--indeed, the likelihood---that some of these paradigms are _mistaken_. There are, after all, many sorts of animals and inanimate objects which, under

the right perceptual conditions, could be mistaken for cats; and such ostensive blunders undoubtedly have been made and gone undetected. The result is either that Final Science provides no significant common classification for these heterogeneous paradigms, in which case no extension for 'cat' is specified, or else some extremely high-level classification is found, in which case the extension of 'cat' becomes promiscuously inclusive, perhaps embracing certain statues, rock-formations, and small dogs! Either result seems intolerable. Just as we appealed to Final Science to anchor one end of the sameness-relation, so too we must find some unequivocal source to anchor the other. It is here that the New Theorist, drawing an analogy with his treatment of proper names, introduces the notion of an historical paradigm, the thing or things first called by the name 'cat' (or its linguistic ancestor). By anchoring the sameness-relation in these *ur*-cats, we can ignore subsequent ostensive misadventures as irrelevant to the determination of extension and can correctly identify them as mistakes. (What has been said thus far about the common noun 'cat' could equally be said for mass terms like 'water', the latter being Putnam's favorite example and the one explicitly discussed by Zemach.)

We are now in a position to formulate a minimal version of the New Theory of Reference for mass and count nouns by claiming that what has been said of 'cat' (and could be said of 'water') holds as well for many other nouns: viz., that their extension is determined, not by associated intensions, but by their connection with historical paradigms at one end and the classifications of Final Science at the other. More precisely, and relativized to English, we may formulate the Minimal Theory as the following series of connected claims:

(MT) Many mass and count nouns of English are such that:

(i) they lack intensions in the traditional sense;
(ii) they nevertheless have determinate extensions as ordinarily used;
(iii) they have these extensions by virtue of being historically descended (via chains of borrowing) from original uses of these words in the context of certain dubbing ceremonies;
(iv) these ceremonies fix the extensions of the words, by reference to an ostensive paradigm, to include all and only that which is of the same kind as the paradigm;
(v) Final Science determines sameness of kind.

In isolation, (MT-ii) and (MT-iii) are not a source of controversy, since the Intensionalist can always think of introductory ceremonies as creating synonyms; so we may expect his scorn to fall upon (MT-i), (Mt-iv), and (MT-v). Let us begin with (MT-iv) and postpone treatment of the others until we have seen what emendations might need to be made.

3. INDEXICALITY AND PARADIGMS

(MT) has been contested even as an account of simple cases like 'cat' and 'water' on the ground that (a):

> ...no one knows, nor can ever hope to know, on what occasions and with respect to what objects our ordinary English substance-terms were first uttered....[Hence] we do not and cannot know what is the correct reference of the word 'water', or any other English substance-term.... (Zemach [1976], p. 123)

and on the ground that (b) given a defective historical paradigm we would be forced to deny something which is "conceptually necessary", viz., 'that our [contemporary] prime example of water is water' (Zemach [1976], p. 124).

Zemach's (b)-criticism depends partly upon an overly literal reading of (MT-iv). If we imagine that a single historical episode of ostensive dubbing infallibly fixes the extension of 'water' for all succeeding generations, then it seems that we would have to admit that the original sample's not being H_2O (but rather some superficially similar substance XYZ) brings it about that all subsequent uses of 'water' refer to XYZ rather than to H_2O, hence that virtually all the stuff we currently call 'water' is not water at all! And this is truly counterintuitive: our untutored verdict would be just the reverse, viz., that our stuff is water but the original sample was not. But why should we affirm the antecedent of this provisionally granted conditional? Even supposing that just one historical dubbing is at issue, why must it be treated as infallible?

Our hypothetical original dubber must not be supposed to be working in a total vacuum. Behind his introduction of the new word 'water' (or its historical antecedent) was a background of experience with water and a battery of referential intensions, both of which are crucial as regards the proper interpretation of just what is being accomplished in the dubbing ceremony. Our dubber is trying to introduce a term to cover the sort of stuff with which he is antecedently familiar (it falls from the sky as rain; it runs in rivers; it has such-and-such sensory properties, etc.). He already has crude ways of grouping together bits of the stuff, and the

demonstrated sample is just a heuristic device to focus his own and the audience's attention. To put it simply: he already _has in mind_ various quantities of stuff which, together with their ilk, he wishes to collect under a term, and he relies on the ostended sample to be more of the same. The paradigm sample thus plays a purely derivative role. The primary question is: what did he have in mind in the first place (which led him to pick _that_ particular sample)? Insofar as this question is to be answered by consideration of what stuff(s) played the dominant causal role in bringing him into the relevant psychological states--and I'm not sure how else it could be answered--the verdict is most likely that it was water, good old H_2O. This is the most reasonable interpretation, in _our_ terms, of what the _ur_-user was trying to denominate.

To this it will be replied, and with some justice, that the problem has been relocated but not eliminated. What if, when all is said and done, it turns out that the dubber lived in a peculiar place, where the locally plentiful liquid he previously experienced was all XYZ, not H_2O, so that what by any reasonable account he "had in mind" was XYZ? Doesn't this doom us to be talking about something other than H_2O when _we_ use 'water' today? Not quite. For we need not follow Zemach in interpreting (MT-iii) as making the strong claim that there is a _single_ historical introduction of 'water' such that every contemporary use of 'water' has its extension fixed by reference to that event. Rather, we can and should interpret (MT-iii) as making the weaker claim that each such contemporary use has its extension fixed by some such introductory event _or other_. Thus we may concede that the "first" use of 'water' and those uses descended from _it_ might be deviant, without having to admit that _our_ uses of 'water' would be deviant. Though necessary in any event, this emendation is still insufficient to deal with Zemach's objection, for we may doggedly raise it again with respect to whatever dubbing(s) our current uses of 'water' _do_ owe allegiance to. Although we have achieved a slightly more perspicuous view of its setting, the problem of the defective dubbing seems to remain.

The shortcoming of the qualifications to (MT) introduced thus far is that they serve only to reduce the _probability_ of contamination from defective dubbings; but Zemach's point pertains to the very _possibility_ of our current uses of 'water' stemming from a defective dubbing--something which can never be ruled out altogether. And this abstract possibility supposedly entails another, viz., that our current paradigms of water might not fall within the extension of 'water' as currently used. But this, Zemach claims, is absurd; and (MT), which allows for this possibility, must be rejected. But _is_ the consequence in question really so absurd? To see that it is not, it will be helpful to consider a speech-community

consisting of just a single individual, so that the complexities of group interaction do not distract us.

Consider young John, who, after learning a little English, was orphaned and left to grow up alone in a desert devoid of water. Luckily, there is an oasis nearby from which he obtains food and drink. But the solitary pool from which he drinks contains a clear, colorless liquid which is (you guessed it!) not H_2O but rather life-sustaining XYZ (left behind by alien visitors). John introduces the word 'water' for himself as the name of the kind of stuff in the pool. A few years later John decides to escape to a more hospitable clime. Filling his canteen from the pool and packing some food, he journeys until he sights the promised land--a valley below with rivers and lakes full of H_2O. Naturally the first thing he says upon sighting the area is 'There is a lot of <u>water</u> there--more than I've ever seen!' If we look just at this <u>initial</u> utterance, made upon first glimpse of the promised land, then there is no absurdity at all in <u>our</u> saying that John is wrong, that 'There is a lot of water there' is a <u>false</u> sentence of John's language relative to the context of utterance. Nor is there any implausibility in supposing that John would <u>admit</u> to being wrong if, at that very moment, he were informed that the stuff he is looking at with such awe is <u>not</u> the same liquid as that in his canteen. However, once John has dwelt in the promised land for a few months or years, matters are far less clear. His primary intension in using 'water' is now to refer to the locally plentiful liquid, and if he were trying to teach his word 'water' to a child, he would point to the local lakes and rivers. Were he <u>now</u> to be informed that the local liquid is different in kind from that in the desert pool of bygone days, it is unlikely that he would change his linguistic ways. Indeed, if enough time had gone by, he might well assent to 'The stuff in the desert pool was not water'. Because of his overriding concern with the <u>local</u> liquid, he behaves just like one who has <u>re</u>introduced 'water' by reference to a new paradigm.[2]

What Zemach overlooks is this tendency on the part of individual speakers tacitly to readjust their usage of certain terms. But there is nothing in (MT) which rules out such readjustments. In fact, it is predictable that such reintroductions will take place for terms such that (i) the associated Stereotype is fairly detailed and (ii) the speech community is preoccupied at different times with different prominent local things or stuff fitting the Stereotype.[3] It is precisely in these cases that the speech community does not defer to history but takes matters into its own hands. On the other hand, where the Stereotype for a term is weak or vacuous and the local things or stuff fitting it are of little or no importance to the community, we find more reliance on historical usage and a corresponding willingness to admit to possible

error in current uses. In neither case does (MT) invite any absurd consequences.

Turning now to Zemach's (a)-criticism, we must ask whether (MT) in fact implies that "... we do not and cannot know what is the correct reference of the word 'water', or any other English substance term." The problem here is to determine just what *counts* as "knowing what the correct reference is". At one level the answer is easy: I *can* know what the correct extension of 'water' is, because I can know (e.g., by authority) that *the extension of 'water' is none other than the totality of stuff which is of the same kind as whatever my ancestors (I myself, my teachers, or whoever) dubbed 'water'*. If the italicized equation is analytic, then (MT) can scarcely be doubted; and if (as I would maintain) it is true but not analytic, then it seems to be a perfectly reasonable and correct answer to the question 'What is the extension of 'water'?' Any objection that I could not come to *know* that the extension of 'water' is just what (MT) says it is would equally seem to impugn the idea that *any* theoretical identity could ever come to be "known". Now if knowing a correct and nontrivial answer to the question ⌜What is the extension of *T*?⌝ does not constitute knowing what is the correct extension of *T*, what on earth would?

Zemach will not accept the proposed answer because it yields no decision procedure. His idea seems to be that we can be certain that such-and-such contemporary quantities of stuff are correctly called 'water', and that this certainty is explicable only on the supposition that we have a decision procedure for the application of 'water' (provided by its "intension"). If (MT) were a correct account of the semantics of 'water', then no such decision procedure would be forthcoming. Hence (MT) must be inadequate.

Given what was said above about the connection between strength of Stereotype, importance of the local stereotypical stuff, and tendency to reintroduce a term, it should be clear what is wrong with Zemach's argument. Briefly, the terms of which we are willing to insist that they "must" correctly apply to such-and-such current candidates are precisely those which, like 'water', have strong Stereotypes and are backed by community interest in local stereotypical stuff to the point of being continually reintroduced by reference to *local* paradigms. We certainly can and do know that these current paradigms are correctly called 'water', for we have fixed the reference of 'water' with respect to *them*. And if we knew enough Chemistry and had the right equipment, we could assay samples of liquid for sameness of kind with these local paradigms. On the other hand, terms like 'elm', which have weak Stereotypes ("a kind of tree") and are not backed by community-wide interests, *do* fit Zemach's picture of being used by reliance on historical dubbing(s) of (possibly long-lost) objects. Yet

these are terms of which we are _not_ willing to assert that they "must" correctly apply to such-and-such local things. (There is no tree in my environment of which I would insist that I am certain that _it_ is an elm if anything is!)

So far we have seen no reason to make any drastic revisions in (MT), though we have found it necessary to add glosses regarding multiple introductions, redubbings, and the derivative character of ostension in reference-fixing. Admittedly, however, these qualifications do undermine the original picture of the unity and continuity of reference. Since we do not insist on eternal reliance upon a single introductory event for each term, the possibility of extension-shifts becomes a real one. Nevertheless, we can still allow for--indeed, predict--unity and continuity of reference within a given speech community over a long period of time, bounded at one end by an initial introduction of a term and at the other by a _de facto_ redubbing under the collective pressure of autonomous referential intensions regarding local stereotypical stuff or things. And within such stretches we can still distinguish "change of belief" from "change of meaning".

(MT) will, however, need substantive revision. For (MT-iv), by requiring ostension of a (would-be) paradigm, thereby restricts the account to nouns the members of whose extensions are (or can be made to be) _observable_. Hence it would appear that _theoretical_ terms, which are typically introduced to refer to posited entities that are not observable at the time, are _ipso facto_ excluded from the scope of the New Theory (cf. Mellor [1977]). And this of course is the exact opposite of what Putnam intended: the New Theory is supposed to provide a model of continuity of reference across theory-changes in science, which it can scarcely do unless it applies to theoretical terms. The needed emendation in (MT) was already implicit in our demotion of ostensive paradigms in favor of "what the dubber had in mind". For to have something in mind is, usually if not always, to have it in mind _under some description or other_. This is not to suggest that the description in question must be "pure" (context-independent); typically it will contain indexical elements. The point is rather that none of these indexical elements need demonstratively refer to any observable paradigm: they may serve only to provide the spatio-temporal framework within which the descriptive elements single out the intended subject. Reference to unobservables thus poses no obstacle in princple, provided that some plausible account can be given of the descriptions allegedly used to "fix" such reference.

The account most frequently mentioned by New Theorists has reference for theoretical terms being fixed by indexical descriptions which invoke underlying causal mechanisms--e.g., 'the cause of these (ostended or described) phenomena', 'the effect (product) of such-and-such processes', etc. As always,

reference-fixing is to be sharply distinguished from synonym-creation. The introducing descriptions do not become synonymous with the terms they are used to introduce. The scenario envisaged earlier in (MT-iv) and (MT-v) thus turns out to be a special case of what we will call the Comprehensive Theory:

(CT) Many mass and count nouns of contemporary (technical and nontechnical) English are such that:

(i) they lack intensions in the traditional sense;
(ii) they have determinate extensions as ordinarily used;
(iii) they have these extensions by virtue of being connected (through chains of borrowing) with earlier uses in which they were introduced into the speech community;
(iv) each such term-introduction overtly or covertly employs an indexical description <u>D</u> such that the extension of the term is fixed as (what Final Science will regard as) the extension of <u>D</u> at the time and place of introduction.

In the simple cases we originally examined --'cat' and 'water'--the introducing description was presumably a relational predicate, the relation being sameness-of-kind and the other term of the relation being the ostended or described paradigm. In the more complicated cases of causal description, <u>D</u> is still relational in form, but the relation and other term[s] have changed.[4] Terms which conform to (CT) may, following Goosens [1977], be called <u>underlying trait terms</u>. (This terminology is preferable to 'substance-term' and 'natural-kind term' inasmuch as it does not suggest that the terms in question must always pick out genuine substances or natural kinds; often they will do so, but (CT) does not prevent the introduction of terms to mark purely arbitrary groupings--e.g., let 'wug' be the common name of all and only those things (whatever they were) seen by me prior to my ninth birthday.)

4. THE ROLE OF "FINAL SCIENCE"

There is a widely popular objection to the New Theorist's account which I have suppressed until now, on the ground that the full resources of (CT) are required to deal with it. The objection comes in several versions, each with a somewhat different emphasis, but a common theme runs throughout: viz., that the deference to <u>science</u> embodied in (CT) yields a grossly distorted picture of how ordinary, <u>non</u>technical mass

and count nouns of English are actually used. In particular, Dupré [1980], McKay and Stern [1979], Mellor [1977], and Zemach [1976] all complain that introduction of a term via a relational predicate like 'is the same kind of thing/stuff as this (ostended or described) thing/stuff' runs afoul of the fact that science may come to regard the paradigmatic things or stuff in question as being of many different kinds, with the result that the New Theorist would have to deny that the term ever had a determinate extension! It is pointed out, e.g., that many familiar chemical elements have isotopes, and that a naturally occuring sample of such an element will be a mixture of these isotopes. Then again, sometimes minerals known by a common name turn out to be chemically quite distinct: jade is not one but two distinct substances, jadite and nephrite. The situation with common names of plants and animals is even worse: at best they rarely pick out anything a botanist or zoologist would count as a species but seem instead to correspond to various higher taxa; and at worst they fall short of corresponding to any taxonomic category, by virtue of lumping together things which the taxonomist regards as importantly different and distinguishing things which he regards as importantly similar. Careful consideration of such examples will lead us to make a number of refinements in (CT).

Talk about sameness of kind is admittedly sloppy. It invites paraphrase by talk of "relevantly similar things", but there is no gain in clarity until some standard(s) of relevance are specified. The putative role of Final Science in (CT) is to act as ultimate arbiter of relevance: the relevant similarities are the ones Final Science uses as the basis for its final classifications. Where isotopes are concerned, this hard line works reasonably well. By the lights of contemporary science (which we may assume to be in step at least to this extent with Final Science), atomic number is taken to be more significant for classificatory purposes than is mass number, with the result that atoms having the same atomic number but different mass numbers are classed together as different "forms" of the same element. Thus the fact that virtually any sample piece of iron is a mixture of Fe^{54}, Fe^{56}, Fe^{57}, and Fe^{58} atoms is perfectly compatible with (CT)'s picture of 'iron' having a determinate reference to the element iron. The isotope-case, far from telling against reliance on Final Science, seems quite in accord with it.

A much better case for the critic is that of 'jade'. For the sake of argument, let us suppose that 'jade' was introduced by people exposed to both jadeite and nephrite but, like most of us,incapable of telling them apart. The use of 'jade' proceeds merrily along until it is discovered that the original or current bits of stuff so-called fall into two very different chemical kinds, for which scientists coin the terms 'jadeite' and 'nephrite'. Clearly there is no single

scientific candidate for "the" mineral in question. We must confess that 'jade' does not pick out a unitary natural kind; rather there are "two kinds of jade". Since the uniqueness-presupposition of the term-introduction is not satisfied, the critic concludes that it follows from (CT) that 'jade', like 'unicorn', is an empty term--a consequence which he rightly labels as absurd.

Underlying trait terms, of which 'jade' is supposed to be an example, are, as the very name suggests, typically introduced on the supposition that the superficially resembling things or bits of stuff in question do share a single common nature. But there is no reason why we must regard this motivating supposition as an indefeasible semantic presupposition of successful term-introduction. When we introduce an underlying trait term, we do not pretend to prescience. Within contextually determined limits, we do not regard the eventual discovery of multiple natures as destructive of our enterprise and would not resist reformulation of our introducing descriptions in terms of "having the nature, or one of the natures, exemplified by such-and-such". Of course our preference is for single underlying natures, and if the discovery of diverse natures closely followed the term-introduction, it is likely that the term, not yet being entrenched in usage, would be withdrawn and new terms introduced by appeal to the newly revealed natures. But this is symptomatic merely of a desire for tidiness of nomenclature; it does not amount to the admission that the original term failed to get any grip on reality. Indeed, the original term, with its admittedly disjunctive extension, is often retained but adjectivally qualified so as to produce compound terms picking out the newly recognized kinds. Thus, e.g., jewelers speak of "true jade" as opposed to "nephrite jade". Clearly, however, this tolerance for multiple natures and correspondingly disjunctive extensions has its limits. Given the various purposes for which an underlying trait term might be introduced, we would expect varying degrees of openness to the proliferation of kinds; but there would inevitably come a point when there were simply too many kinds of things or stuff in question--so many that the purposes for which the term was introduced could no longer be served. Suppose that, shortly after the introduction of 'cat', people began for the first time to dissect and examine the local things so-called, discovering that no two were internally alike. The first contained cogs and wheels; the second contained a family of mice who operated the "cat"-body like a puppet; the third was completely hollow and inhabited by a dimly visible ghost; and so on through hundreds of cases. Under the pressure of these revelations, one of two things is likely to happen: either the new word 'cat' would be abandoned or else it would be reintroduced as a "gestalt" term, a designation of things that have the external appearance of the

items in question.

Behind the introduction of an underlying trait term like 'jade' are two sorts of desires: on the one hand, the desire to have a term covering the local things whose superficial resemblance in certain respects has made an impression on us (i.e., whose similarities count as "important" relative to our purposes and interests); and, on the other hand, the desire that the term thus introduced should ultimately prove useful in formulating "significant" inductive generalizations--a project whose prospects for success vary inversely with the number of distinct natural kinds found within the term's extension to the extent that "significance" is determined by purely scientific interests. Relative to other interests, however, the proliferation of scientifically recognized kinds within the paradigm-class may not destroy the possibility of using the term to frame many "significant" generalizations. If one is concerned with explanation and prediction only at a superficial level, the deeper distinctions may not make any practical difference to one's enterprise; for even things with diverse natures may turn out to be predictably similar in the respects which are important at the moment. Before pursuing this line of thought any further, however, we must note an additional complication.

We have simply taken for granted, as in the case of 'jade', that Final Science will provide a determinate answer to the question of how many kinds of things or stuff are exemplified in the paradigm class, so that the term in question will be said to have a determinate extension embracing just those kinds. However plausible this supposition may be for cases in which only chemical structure is at issue, it becomes much less plausible when we begin to consider biological kinds. Given a collection of superfically similar organisms, there are many ways of individuating the "kinds" they exemplify, depending on what taxonomical level one has in mind: species, genus, order, etc. If a term like 'beetle' or 'lily' is to have a determinate extension, there must be a determinate principle for sorting the paradigm-class; for it will make a vast difference to the extension whether we include in it simply the members of the particular species present in the sample or include also the members of higher taxa exemplified by those species. One of the points made by Dupré [1980] is that there seems to be no nonarbitrary way of settling this matter. (Dupré supposes that a single taxon must be at issue, but the problem obviously remains even if we allow for many-sorted extensions.) No matter what taxonomical level is picked, there will be trouble somewhere. If, e.g., we count only the species present in the sample, then we stipulate an absurdly narrow extension for words like 'beetle'; for there are hundreds of thousands of species of beetle, of which only a small number are likely to have been represented in any

actual sample. If. e.g., we insist on the "lowest" taxon embracing all the individuals in the sample, we produce absurdly broad extensions for other terms and end (to use Dupré's example) by counting onions as lilies and butterflies as moths.

We have seen that the isotope-case and the jade-case can be reconciled with the paradigm-and-sameness-relation model inherited from (MT). The case of biological kind terms, however, appears to require a different treatment. Since (MT) was extended to (CT) to allow for the treatment of a broader class of terms--in particular, theoretical terms--we must look to the richer resources of (CT) to handle the problem of "common names" of plants and animals.

Suppose we begin with a class of creatures demarcated by "looks". Eventually, we notice that things which look alike in the relevant respects nonetheless differ with respect to various other contextually interesting properties not included in the specification of their looks. And other differences are found to be predictably related to these, until we have clusters of interesting properties that divide our original group into subgroups. At some point we become sophisticated enough to suspect that there is "something behind" each of these reliably related clusters of features, that if only we knew enough, we could explain the differences between these subgroups. The time is now ripe for the introduction of underlying trait terms by means of descriptions of the form 'things having the underlying nature(s) responsible for such-and-such a cluster of interesting properties being exhibited by certain things of such-and-such _gestalt_'. Different interests will pick out different clusters of properties for the underlying nature(s) to tie together. By the lights of Final Science, there may be a number of distinct structural traits possessed by things of the relevant _gestalt_ which would explain the connections between the important properties in question. But--and this is the crucial point--there is _no_ reason why these traits need answer to anything of fundamental scientific interest; in particular, there is no reason why they must turn out to be identical with the traits used by (present or future) biologists as the basis for their taxonomy of organisms (cf. Ware [1978], pp. 47-8).

We can, e.g., imagine a forest-dwelling tribe whose interest in the surrounding trees is purely that of woodworkers. Within the tree-_gestalt_, "trees" are classified by the properties of their wood which are most relevant to woodworking: texture, grain, hardness, suppleness, etc. Many of these properties are dispositional in nature and are reliably related to each other and to various observational properties of wood. We can readily imagine the tribe supposing that there is "something deeper" which explains why these properties form inductive clusters, and going on to introduce

underlying trait terms to name the "kinds" of trees in their forest, the extension of each name being fixed to comprise those "tree-ish" things which have the underlying structure(s) responsible for yielding wood with such-and-such a cluster of important, inductively related properties. But such a system of "common names" for "kinds" of trees is unlikely to correspond very closely with a botanist's taxonomical survey of the forest. Final Science is still relevant for ultimately identifying underlying traits and determining the extension of the tribe's terms, but the purposes for which these terms were introduced can infect the introducing descriptions in such wise that the resulting extension-by-the-lights-of-Final-Science does not match any taxonomical category deemed important by Final Science itself. The critics have misunderstood (CT) if they think that reliance on Final Science somehow entails that our ordinary underlying trait terms for "kinds" of plants and animals must _converge_ with the taxonomical designations of present or future biology.[5]

According to this picture, ordinary underlying trait terms for "kinds" of plants and animals are typically parasitic on related _gestalt_-terms and a certain amount of antecedent interest-relative sorting of items within the _gestalt_. Given the historical, community-wide importance of various plants and animals (e.g., as food-sources) it is not surprising that there is some convergence between "common" and scientific classifications (especially at the level of species of large mammals). For our mundane concerns often overlap those of the scientist. A farming community has a deep and abiding interest in, e.g., the _reproductive_ characteristics of the plants and animals they breed and crossbreed--a point of obvious contact with the biologist's phylogenetic concerns. At the opposite extreme lie plants and animals in which there is little or no community-wide interest at all, hence no motivation for the interest-relative sorting within the initial _gestalt_ which might lead to the introduction of an underlying trait term, or to the reintroduction of the original _gestalt_-term as an underlying trait term. Here a _gestalt_-term is all that is needed.

Of course, human curiosity being what it is, there may arise groups of nature-lovers who become experts with respect to the very creatures which the community as a whole tends to ignore. These experts, after engaging in some morphological sorting, might then introduce corresponding underlying trait terms. But unless these naturalists become "acknowledged" by the community, so that ordinary speakers defer to their judgement by borrowing from them, the underlying trait terms they introduce will never become serious competitors to the _gestalt_-terms employed by everyone else. Terms like 'beetle' and 'lily', I suspect, survive for most of us as simple _gestalt_-terms, though the progress of scientific education may soon

alter this fact. The upshot for (CT) is mixed. On the one hand, we have seen reason to think that many ordinary biological kind terms can plausibly be construed as underlying trait terms, though only a small number will have extensions corresponding to "biological kinds" in the sense of scientifically recognized taxa. On the other hand, we have been forced to concede that a good many common names of plants and animals probably have not yet achieved the status of underlying trait terms for the majority of English-speakers. To the extent, however, that gestalt-terms tend to evolve into underlying trait terms (at least for certain special-interest groups within the speech community), the importance of (CT) will be undiminished.

5. UNDERLYING TRAITS AND ANALYTICAL DEFINITIONS

The New Theorist's talk of "underlying traits", "hidden natures", and the like has generated some confusion anent the claim (Putnam [1975a], pp. 243-244) that even artifact-terms are, or can evolve into, underlying trait terms. Schwartz [1978] takes Putnam to task on this count, arguing that members of artifactual kinds do not have, nor are they commonly presumed to have, common underlying natures:

> What makes something a pencil are superficial characteristics such as a certain form and function. There is nothing underlying about these features. They are analytically associated with the term 'pencil', not disclosed by scientific investigation. (Schwartz [1978], p. 571)

Schwartz asks us to imagine that pencils are discovered to be organisms of some sort. Still, he insists, we would call any subsequently manufactured item having the right form and function a pencil, regardless of the fact that it did not share the organic nature of the "old" pencils. But, as Kornblith [1980] correctly points out, this argument clearly rests upon a false assumption: viz., that "underlying nature" can only be understood in terms of a thing's composition. Once this error is discarded, Kornblith says, there is nothing to prevent the New Theorist from saying that function is the underlying nature of most if not all artifactual kinds. Whatever one may think of this positive claim, this much seems clear: underlying natures, though often "hidden" in the sense of being unobservable or unknown, need not be so. Kripke [1972] provides a case in point with his example of the Standard Meter Bar. We may fix the extension of 'meter' by the indexical description 'length equal to the distance between the two marks on this bar', thus making 'meter' an underlying trait term in our sense; yet nothing mysterious, unobservable, or

unknown seems to be involved here.

When the function of an artifact is not known, there is little or no intuitive resistance to the idea that our name for the relevant kind may be an underlying trait term. Kornblith cites the example of an archaeologist who digs up a collection of similar objects which are obviously implements of some sort but which are so unlike familiar implements that he chooses to call them by a new name, whose extension he fixes by the description 'implements having whatever function(s) these things were designed to serve'. Of course this is not the standard case. Typically, artifacts are created to fulfill an antecedently specified function and are named by the reference to just that function, which is usually explained to us when we are taught the name. Because everything relevant is "out in the open", there is a strong inclination to think that such terms must be synonymous with the corresponding functional descriptions. This conviction is extremely difficult to shake, even when one recognizes the distinction between reference-fixing and synonym-creation. In the parallel case of the Standard Meter Bar, we may be pursuaded to opt for reference-fixing and to deny the synonymy on the ground that, e.g., it makes sense to suppose that the marks on the bar could have been more or less than one meter apart. But similar attempts to pry artifacts apart from their functions are less convincing. So let us approach the matter from a different angle than Kornblith does.

The first thing to notice is that underlying trait terms can be used in descriptions to produce new terms which are synonymous with those descriptions. Given, say, the previously introduced underlying trait term 'cat', we may invent a new term, 'kitten', and define it by 'juvenile cat'. Now 'kitten' is not, strictly speaking, itself an underlying trait term; yet, owing to the essential presence in its definition of an underlying trait term, 'kitten' is not a simple "nominal kind" term like 'bachelor' either. Such hybrid terms serve one of two functions: they may serve, like 'kitten', to restrict the extension of some underlying trait term to a proper subset; or, like 'bullfighter', they may pick out a new class of things defined by the relation of those things to items in the extension of an underlying trait term. Let us call these two sorts of hybrids "trait-restrictive" and "trait-parasitic" respectively. The crucial point about such hybrids is that, although (CT) does not apply to them directly, we ultimately need (CT) to explain their workings. I want to suggest that something similar may be true of many artifact terms, hence that (CT) may ultimately be needed to account for them, even though they are not underlying trait terms in the pure sense.

Actually, it is quite easy to show that at least some artifact terms are trait-restrictive or trait-parasitic

hybrids. We need look no further than those generic artifact terms which are defined by reference to composition as well as to function. For the specification of composition usually employs an underlying trait term for the relevant "stuff". Examples are 'silverware', 'glassware', 'ironware', 'woodcut', 'china', etc. Even when composition is not essential, underlying trait terms are often employed in the specification of function. Thus, 'sphygmomanometer' is defined in standard dictionaries by the description 'device for measuring arterial blood pressure', where 'blood' is an excellent candidate for being an underlying trait term. 'Coffeepot' is defined by 'container for brewing and/or serving coffee', and so on through a fairly large number of cases. Occasionally, both composition and function are specified via underlying trait terms--e.g., 'cold chisel' is defined by 'hardened and tempered steel chisel without a handle, for cutting and chipping cold metal'. Finally, underlying trait terms are sometimes used in the specification of the form of an artifact. 'Statue' is defined by 'carved, cast, or molded representation of human or animal form.' For 'fly' (in fisherman's terminology) the dictionary provides 'device made to resemble an insect and used as bait'. Other examples are 'weathercock' and 'scarab'.

One misunderstanding about these examples must be headed off immediately. There is admittedly a sense in which underlying trait terms could be eliminated from at least some of these definitions: e.g., a purely mechanical description of a sphygmomanometer might manage to pick out just the same class of devices; and a purely physical description of coffeepots might likewise manage to circumscribe just the right class of artifacts. But even if such extension-preserving redefinitions were always possible, it would not show that the expressions in question are not hybrids. For it remains the case that these terms were originally defined and subsequently taught by means of descriptions which do contain underlying trait terms, and that is all it takes for these artifact terms, used in such a tradition, to be hybrids. (Since dictionaries make free use of underlying trait terms in function-specifications, anyone who learns an artifact term from such a dictionary definition will be using it as a hybrid expression.) The possibility of extension-preserving redefinition is no more relevant here than it is in the case of pure underlying trait terms; if we were omniscient, we wouldn't need underlying trait terms--e.g., 'cat' might be analytically defined by reference to some complicated description of genetic structure--but this scarcely shows that we don't have and need such terms in our language.

Moreover, it is not obvious that the envisaged redefinitions could be given in all cases involving hybrid terms for artifactual kinds. 'Coffeepot' is a nice example of the problem. No matter how narrowly we circumscribe the form (and

even the composition), the most we can accomplish is a narrowing of the range of functions that things having that form (and composition) are suitable for. But we cannot thereby capture the decisive factor--viz., which of these possible functions people use the artifacts in question to serve. There are many coffeepots which are indistinguishable in form and composition from teapots, chocolate servers, and water pitchers, all of which might have the same range of potential uses. The coffeepots are such precisely because we use them to brew and/or serve coffee (or stipulate that they are to be so used), and this fact is faithfully reflected in the dictionary definitions.

If one looks closely at dictionary definitions and at the descriptions used to teach artifact terms to children, underlying trait terms crop up again and again. There is no escaping the conclusion that, as actually employed by English speakers, many artifact terms are genuine hybrids. Nevertheless, there will still be plenty of "pure" (nonhybrid) artifact terms. While the descriptions associated with these terms might, as Kornblith suggests, have been used merely to fix their reference rather than their meaning, I can find no convincing argument to show that they were introduced in this way, let alone that they "must" have been. Putnam's claims about artifact terms may have been overoptimistic, but we have at least seen reason to think that (CT) is not completely irrelevant to the matter.

6. THEORETICAL TERMS AS UNDERLYING TRAIT TERMS

As yet we have said little about the application of (CT) to theoretical terms in science, beyond noting that (CT) is in principle so applicable. But while no one would deny that theoretical terms might be underlying trait terms, there is no lack of critics who deny that they actually do function in this way. Thus Fine [1975] complains that

> ... the commonplace idea that the subject of scientific endeavor varies from time to time ... undermines Putnam's account of reference. ... [O]ne can ... choose a subject of study and later change one's mind. (24)

He illustrates this complaint with the history of the term 'electron'. 'Electron' was first introduced in 1891 by G. Johnstone Stoney as a name of the fundamental unit of electricity (i.e., the unit quantity of positive or negative electrical charge), the paradigm being the electrical charge on the hydrogen ion in electrolysis. Under the influence of "particulate" accounts of electricity, the term 'electron' was naturally assimilated to J.J. Thompson's "corpuscles" and used

to refer to the particles that carry the unit charge of negative electricity. Fine continues,

> here ... we have a concept evolving, à la Kuhn, so as to bring with its evolution a shift in reference. ... Nor are their two words, one introduced to refer rigidly in the Stoney way and the other introduced to refer rigidly in our way. For neither of these words fits the true formula, "Once 'electron' referred to a quantity of charge, but now 'electron' refers to a particle." ... [T]his change ... shows the natural evolution of the subject matter of the science of electricity. By attaching the referent rigidly to the word, however, Putnam's account forecloses this very possibility. (26)

That the extension of 'electron' has changed seems undeniable, but it is not clear that (CT) is impotent to account for this fact. Let us suppose that Stoney simply fixed the reference of 'electron' via the description 'positive or negative charge equal to the charge on the hydrogen ion in electrolysis'. In contemporary terms, we would say that Stoney used 'electron' to refer to any charge of $\pm 1.602 \times 10^{-19}$ coulomb (call this charge '*e*'), and uses borrowed from his had the same reference. At some later point it was decided, quite naturally, that 'electron' would be a handy and appropriate designation for non-nuclear particles of charge -*e*. But this decision was surely not made in ignorance of the fact that 'electron' already had a history; indeed, it is precisely *because* of the recognized historical connection of 'electron' with *e* that it is so natural to reapply the term to particles of charge -*e*. In other words, there was a conscious (and well motivated) *break* with Stoney's tradition. 'Electron' was *re*introduced as a particle-name and a whole new tradition of usage initiated. Fine seems to think that, on Putnam's view, once a term's reference is historically fixed, it cannot subsequently come apart from that reference. But (CT) implies continuity of reference only *within* an historical tradition of cooperation and borrowing; there is nothing in (CT) which rules out discontinuities due to *re*dubbings and subsequent, differently based, traditions of usage.

In the second of the quoted passages above, Fine denies that any such redubbing ever took place, but the reason he gives (his "true formula"-test) is very puzzling. Apparently, he has in mind some argument like the following. If we say that 'electron' was reintroduced, then we in effect are saying that there are two distinct but homophonous words, 'electron_1' and 'electron_2', the former being Stoney's word and the latter being ours. But the true and univocal formula "Once

'electron' referred to a quantity of charge, but now 'electron' refers to a particle" becomes false if either '$electron_1$' or '$electron_2$' is uniformly substituted for 'electron'. So this univocal truth cannot be accounted for on Putnam's view. This argument, however, is seriously confused. First of all, it is word-tokens which have reference, not word-types; this is perfectly obvious for singular terms and is no less true for general terms. To make sense, Fine's formula must be recast as something like

(F) Once (all) tokens of the word-type 'electron' referred to a quantity of charge, but now (most) tokens of that type refer to a kind of particle.

But (F) is a straightforward consequence of the Putnamesque account of 'electron': early tokens of the word-type 'electron' (which stemmed from Stoney's dubbing) did refer to a quantity of charge, and most current tokens of the word-type 'electron' (viz., all those stemming from redubbing) refer to a kind of particle. The "two words" ploy is a red herring. By failing to distinguish type-considerations from token-considerations, Fine invites us to read his formula in wide-scope fashion as

(F_1) 'Electron' is such that once it referred to a quantity of charge, but now it refers to a particle.

and then he notes that neither '$electron_1$' nor '$electron_2$' can replace 'electron' *salva veritate*. But this is irrelevant, since the sense in which (F_1) is true is just that in which it is equivalent to

(F_2) The word-type 'electron' is such that once its tokens (=those produced at an earlier time) referred to a quantity of charge but now its tokens (=those produced currently) refer to a particle.

and not equivalent to

(F_3) The tokens of 'electron' are such that once they referred to a quantity of charge but now they refer to particles.

And (F_2) is equivalent in turn to (F). There is thus nothing in Fine's illustration that need worry the adherent of (CT).

A much deeper objection--but one still similar in spirit to Fine's--has been developed by Enc [1976]. Enc attempts to

show that the plausibility of (CT) varies directly with the "observationality" of a term, hence that (CT), however successful at lower levels, cannot begin to cope with high-level theoretical terms. He begins by noting a striking asymmetry between an ordinary, observation-based term like 'cat' and its low-level theoretical counterpart '_Felis domesticus_'. If we began for the first time to dissect the things we had called 'cats' and discovered nothing but machinery inside, then we would (as Putnam predicts) simply say that cats had turned out to be robots, not animals (as previously thought). What we would _not_ say is that there aren't any cats. But the biologist would, using his counterpart-term '_Felis domesticus_', say just the opposite: viz., that the species _Felis domesticus_ doesn't exist. We would not expect him to say, 'The _Felis domesticus_ turns out not to be an animal'. This story supposedly illustrates a common phenomenon in science: the scientist will deny that _F_s exist rather than admit that they are radically different from what his theory says they are. Consider the case of phlogiston-theory. If we take 'phlogiston' to have been introduced as an underlying trait term to name the substance responsible for certain phenomena such as combustion and calcination, then, since _oxygen_ is in fact the substance responsible for these phenomena, (CT) would dictate that 'phlogiston' referred to oxygen, about which phlogiston-theorists went on to form a number of unfortunately mistaken beliefs. But, Enc urges,

> ... Lavoisier, in discovering oxygen, discovered that phlogiston does not exist. ... [T]he phlogiston theorists ... were not talking about oxygen. They were not as a matter of fact talking about anything that exists. Even if Lavoisier had kept on using the word "phlogiston" for his substance, he still would not have been talking about the same substance that the phlogiston theorists were talking about. ...[T]he fact that a theoretical term is retained when the theory undergoes important changes is not by itself a guarantee that when we use the term after the change we will still be talking about the same thing (or things). In fact our judgment as to whether we are talking about the same thing seems to determine whether we should retain the theoretical term in the revised theory. (Enc [1976], p. 268)

With terms which are ostensively based (or capable of being ostensively based), we already know in part _which_ things we are talking about, and we go on to discover their essences (kinds). But with higher-level theoretical terms, which are

not so based, matters are reversed: we cannot know which things we are talking about until we know _what kinds_ of things they are supposed to be.

Not surprisingly, Enc's picture of how (higher-level) theoretical terms _do_ acquire their reference is a version of "implicit definition" Intensionalism. The idea is that investigation of phenomena leads to the hypothesis that a unique non-ostensible object or kind of object is causally involved in their production, whereupon various properties are attributed to the hypothetical object(s) in question, and various mechanisms are posited to connect possession of those properties with the production of the phenomena. If the resulting explanation is successful, these properties will be thought of as constituting a real _kind_; and if the scientist believes that it is a _new_ kind, he will coin a name for it. The reference of this name is determined by these kind-constituting properties and explanatory mechanisms, which are analytically associated with the name and constitute its intension. By contrast, Enc claims that Putnam's view is both false as a description of the actual history of many theoretical terms and hopeless as a normative proposal for the rational reconstruction of scientific activity. It is unreasonable as a normative proposal because the whole idea that a theoretical term is introduced to refer just to whatever is responsible for such-and-such phenomena runs counter to the very considerations which induce scientists to introduce new theoretical terms in the first place. Introduction of a new theoretical term presupposes belief that a new (kind of) object is involved, which in turn presupposes some notion of _what_ kind of thing it is (in contrast to familiar kinds)--and that notion is provided by the properties and mechanisms posited by the containing theory. "Whatever causes such-and-such" cannot constitute the notion of a _new_ kind which contrasts with familiar kinds: such a specification is neutral in just the wrong way!

Let us consider Enc's points in the order given. First, we may agree completely with the raw data concerning the asymmetry between 'cat' and '_Felis domesticus_', but we need not accept Enc's explanation of those data. The species term '_Felis domesticus_' is part of an elaborate scientific scheme for classifying _organisms_ (living creatures composed of cells), and of course such a term would be withdrawn as idle if things of the appropriate phenotype universally turned out to be robots. One explanation of this fact, suggested by Enc, is that species terms are analytically defined via the phenotypic and/or genotypic factors which form the theoretical basis of Taxonomy; the discovery that cats are robots would leave '_Felis domesticus_', like 'unicorn', a meaningful but empty term, which would consequently be discarded as of no practical value to science. But this is not the only picture

which fits the data. The relation of phenotype to genotype can be thought of as analogous to that between Stereotype And underlying trait(s), the former relation being a vastly more sophisticated counterpart to the latter. Accordingly, species terms could be thought of as underlying genetic trait terms introduced by means of phenotypic descriptions (plus, perhaps, ostension of a type-specimen exemplifying the phenotype). Thus e.g., '*Felis domesticus*' could have been introduced by a reference-fixing description of the form 'organism having the genetic constitution of these (phenotypically-*cum*-demonstratively specified) things'. Under the cat-robot hypothesis, such a term-introduction would be a *misfire*: the described and/or ostended things are not organisms and have no genetic constitution. So '*Felis domesticus*' is once more judged to be an empty term and discarded.

To this alternative explanation it might be replied that, if successful, it would undercut our ability to account for the initially noted asymmetry between '*Felis domesticus*' and 'cat'. For, by the lights of the New Theory of Reference, 'cat' is *also* supposed to be an underlying trait term. How is it that we can retain one and reject the other in the cat-robot case? The answer is that 'cat' was introduced in a very liberal, nontechnical way via some description like 'things of the same kind as *these* (ostended or described) objects'--where "being of the same kind" may be understood, in light of our earlier discussion of common names of plants and animals, in terms of having whatever underlying nature accounts for the inductive cluster of contextually interesting properties antecedently found to be associated with items fitting the cat-*gestalt*. The original introducer of 'cat' (or its linguistic ancestor) was presumably innocent of such technical notions as "cellular organism", "genetic constitution", and the like. As it (hypothetically) turns out, cat-robots *do* form a kind--an artifactual kind like clocks. The original reference-fixing ceremony is thus in no way compromised. Apprized of the relevant facts, we would continue to use 'cat' to refer to those objects, though of course we might have to give up some previously accumulated beliefs about them. (Notice that it is essential that the things in question *all* turn out to be robots; as we saw earlier, if the things turned out to be of many different and unrelated kinds, then the term 'cat' might well be abandoned or reintroduced as a mere *gestalt*-term.)

What we have said about the cat-robot case has obvious implications for the treatment of the phlogiston/oxygen case, though the latter is considerably more complicated than the former in virtue of involving two full-blown theories containing the terms in question. As before, we can accept the data but dispute the explanation Enc offers. Ironically, Putnam is in complete agreement with Enc as regards the emptiness of

'phlogiston' (cf. Putnam [1973], p. 206: '... a physical magnitude term such as 'electricity' is sharply distinguished from a term for a fictitious or nonexistent physical magnitude or substance such as 'phlogiston'.') Evidently, then, Putnam doesn't think that his view commits him to the alleged consequence that 'phlogiston' refers to oxygen. Why is this?

Presumably Putnam and other New Theorists would want to maintain that both 'phlogiston' and 'oxygen' should be regarded as having been introduced as underlying trait terms. This narrows our options to a single strategy: 'phlogiston' and 'oxygen' cannot be regarded as having been introduced by the same or equivalent descriptions. So we must find some plausible account of the difference between their respective introductions which shows why one misfires whereas the other one succeeds. I shall assume that all is well with 'oxygen' and concentrate on what went wrong with 'phlogiston'. Enc would have us believe that G.E. Stahl, the 18th Century Swedish chemist who invented the term 'phlogiston', could have been using it as an underlying trait term <u>only if</u> he had thought of phlogiston simply as "whatever is responsible for such-and-such effects" (e.g., combustion and calcination). But (CT) imposes no such constraint on underlying trait terms: their introducing descriptions may be as rich and theory-laden as the dubber wishes. In particular, such introducing descriptions may contain some specification of the <u>mechanisms</u> through which the posited underlying trait supposedly operates to produce the mentioned effects (cf. Putnam [1973], p. 200). Indeed, Enc's own description of how new theoretical terms are standardly introduced would lead us to expect this. The historical evidence in fact suggests that phlogiston was initially conceived to be "responsible for" combustion and calcination <u>in the sense</u> of being an element ingredient in objects capable of undergoing these processes and whose "escape" from such objects constitutes their undergoing these processes. Dampier [1929] put it succinctly thus:

> When bodies are burnt, it seems that <u>something escapes</u>. <u>This something</u>, for long identified with sulphur, was called "phlogiston", the principle of fire, by G.E. Stahl. ... (p. 197; my emphasis)

Of course our modern understanding of the reciprocal character of oxidation-reduction reactions shows that there is no such substance, hence that Stahl's introduction of 'phlogiston' was defective and produced an empty term.

Enc fails to appreciate the possibility of <u>hierarchical</u> introduction of underlying trait terms, where ostensively based underlying trait terms are used in theoretical descriptions to fix the reference of new underlying trait terms of

higher level, which may in turn be so used to produce underlying trait terms of yet higher level. As Putnam stresses, high-level theoretical terms may be introduced via descriptions employing antecedently introduced underlying trait terms and various laws connecting them. (CT) in no ways requires, as Enc seems to suppose, that theoretical terms must be introduced, in _non_theoretical vocabulary, by _direct_ reference to the observable data that the containing theory is designed to explain. When this Straw Man is cast aside, we can see that there is no incompatiblity between (CT) and the sorts of considerations which, according to Enc, lead scientists to introduce new theoretical terms. For the posited distinctive properties and mechanisms (which may be described in terms of antecedently introduced underlying trait terms) can be _built into_ the description which is used to fix the reference of the term for the "new kind". In the end, then, Enc's complaint reduces to the very complaint we encountered in discussing artifact-terms: viz., that when the introducing description is sufficiently rich, so that nothing "hidden" seems to be involved, intuition is often on the side of the verdict that the term becomes _synonymous_ with the description in question. Having grappled with this issue earlier, there is no need to recapitulate here. It is worth noting, however, that in the scientific case the hierarchical stratification of underlying trait terms can serve to mask trait-dependence at higher levels. Final Science, if such there be, would by definition be in a position to expunge any reliance on underlying traits; but we have a long way yet to go.

7. RIGIDITY AND ESSENCE

In speaking of underlying trait terms thus far, we have made free use of such notions as their "reference" and "extension". The concerns of this section will force us to be a bit more perspicuous in our terminology. In particular, we will want to distinguish the role of underlying trait terms in subject position from their role in predicate position. In predicate position (as in '... is water' or '... is a tiger') these terms form logically complex predicate-expressions. When we speak of the _extension_ of such a mass or count noun _N_ (at an index _i_) we are henceforth to be understood as speaking of the _set_ of items which satisfy (at _i_) the appropriate predicate-expression ('... is _N_', in case _N_ is a mass noun; '... is a(n) _N_', in case _N_ is a count noun). In subject position, however, mass nouns and pluralized count nouns seem to behave very much like proper names, i.e., as _singular referring_ expressions. Now it is tempting to reason as follows. Since the extension of a mass or count noun at an index is a well-defined object, we can simply say that what a noun in subject position refers to is its extension at the point of utterance (e.g., our

world), or perhaps the union of its extensions at various worlds. But either decision would have unwelcome consequences. If 'cats' referred in our mouths to the set of actual cats or to the set of possible cats, we could make no clear sense of such true counterfactual claims as '<u>Cats</u> might have been more or less numerous than they are'. For sets cannot change their membership while retaining their identity. Insofar as mass nouns and pluralized count nouns can behave as singular referring expressions, their "referents" must be items distinct from their actual extensions and the union of their possible extensions. It is an open question just <u>what</u> they refer to in such uses.

This open question becomes crucial when we consider the claim, made by Kripke and endorsed by Putnam, that underlying trait terms in subject position not only refer but do so <u>rigidly</u>, picking out "the same thing" in all possible worlds in which "that thing" exists. Until we know <u>what</u> such terms refer to, we cannot understand, let alone assess, the claim that their referents remain constant. Even if we could make sense of the idea that 'cats' refers (at a world <u>w</u>) to the extension of 'cat' at <u>w</u>, this would not help us with the rigidity thesis. For the extension of 'cat' will <u>vary</u> from world to world: there will be no single set which is picked out at each world. Perhaps, then, what 'cats' refers to is cathood or cat-kind. As an abstract object, cathood stays conveniently fixed from world to world, giving substance to the claim that 'cats' is a rigid designator. (Under certain assumptions, the same could be said for the set of all possible cats, but we have already seen that it won't do for other reasons.) But virtually every common noun has, or can be construed as having, a similar property-referring use. 'Bachelors' can be used abstractly to refer to bachelorhood or bachelor-kind, and in such uses it too would designate rigidly. Considerations like these have moved some authors (e.g., Donnellan [1973] and Schwartz [1977], pp. 37-9) to conclude that Kripke and Putnam face a dilemma: either underlying trait terms in subject position are not rigid designators at all, or else the rigid/nonrigid distinction applies equally to all common nouns and so fails to mark any distinctive feature of underlying trait terms.

Two important and easily overlooked facets of the rigidity thesis must be acknowledged at the outset. First, rigidity cannot by itself be used to distinguish between terms to which the New Theory applies and those to which it does not. As Kripke points out, descriptions couched in terms of "essential" properties (e.g., 'the sum of 2 and 5') will pick out the same item in each world containing it, and the same is obviously true of any name analytically <u>defined</u> by such a description. So the class of rigid designators is already admitted to extend beyond the class of terms to which the New

Theory of Reference applies. Second, the rigidity thesis is _not_ a logical consequence of the New Theory _qua_ theory of how the real-world reference of certain terms is determined; rather, it is supposed to follow from the New Theory _plus_ certain substantive proposals about the nature of modal reasoning (i.e., Kripke's "counterfactual" construal of possible-worlds talk). These twin facts suggest that we might want, for present purposes, simply to grasp the second horn of the putative dilemma. Before doing so, however, let us look at an instructive recent attempt at avoiding the problem altogether.

Cook [1980] proposes to escape the dilemma by observing that

> ... 'cat', while not designating the same objects in all possible worlds, does designate the same objects _wherever these objects exist_. What actually exists is world-relative; so the designation of both 'Nixon' and 'cat' is world-relative. But apart from questions of existence, neither varies its designation: a cat cannot exist without being a cat any more than Nixon can exist without being Nixon. (p. 63) ... the only way in which this extension [of 'cat'] varies is by different _cats_ existing in different worlds. (p. 64)
>
> In contrast, 'bachelor' varies its designation apart from questions of existence. Not only can the population of bachelors differ from world to world, those individuals who are bachelors in one world can be non-bachelors in another world. ... Hence 'bachelor' is non-rigid (p. 63).

This way of looking at the matter provides no answer to our question about the singular reference of underlying trait terms in subject position: to say that 'cats' designates _cats_ is to fall back upon an unanalysed notion of "collective reference". Indeed, Cook seems to treat designation-talk as mere shorthand for extension-talk, the latter being his primary concern. To that extent, he is tacitly abandoning the idea that underlying trait terms can actually function as singular terms. This, of course, is not a crushing objection: even if the notion of rigid designation for these terms turns out to be a roundabout way of talking about their actual and possible extensions, it is still worthwhile seeing what "rigidity" comes to.

If we generalize upon Cook's remarks, we obtain the following definition:

(D) For any mass or count noun _N_, _N_ is a rigid designator iff

(i) the extension of _N_ at our world is nonempty (i.e., _N_ "designates"), and
(ii) for any possible world _w_ and object _x_, if _x_ is in the extension of _N_ at _w_, then _x_ is in the extension of _N_ at each world _v_ such that _x_ exists in _v_ (i.e., _N_ is "rigid").

If we accept (D), then we cannot countenance the existence of any rigid designators at all without logically committing ourselves to a strong form of Essentialism. If, e.g., 'cat' is a rigid designator in the sense of (D), then every (actual or possible) cat is _necessarily_ a cat. Because Kripke, Putnam, and other New Theorists are avowedly Essentialists of this stripe, Cook regards (D) as faithfully reflecting what they must have in mind when they claim that certain nouns rigidly designate and others do not. Essentialism about kind-membership is the price we pay for avoiding the dilemma in Cook's way. But if we are going to pay the price, we should at least get our money's worth in return. Cook's account, however, does not provide us with a notion of rigid designation which will do the jobs Kripke requires. For example, equations whose terms designate the same thing are true; and if each term _rigidly_ designates that thing, the equation is held to be a _necessary_ truth--this was supposed to be a striking fact about 'Cicero = Tully' which carries over to the likes of 'Water = H_2O' and 'woodchucks = groundhogs'. Since Cook reduces talk about reference/designation to talk about extension-at-a-world, it is not clear what he can say about the last two equations except to deny that they have the logical form of identities and to paraphrase them by '$\Box\forall \underline{x}(\underline{x}$ is water $\equiv \underline{x}$ is $H_2O)$' and '$\Box\forall \underline{x}(\underline{x}$ is a woodchuck $\equiv \underline{x}$ is a groundhog)' respectively. In place of the remark that these terms refer to the same thing, we would presumably have the claim that their extensions are nonempty at, and coincide at, the actual world. And finally, each term has its extension rigidly in the sense of (D-ii). The problem is that these considerations do _not_ entail the truth--let alone the necessary truth--of the modal biconditionals in question.

A natural reaction to these difficulties is to say that mass nouns and pluralized count nouns in subject position make singular reference to _kinds_ and that kinds are entities distinct from extensions and their members. But kinds, it will be protested, are creatures of darkness whose identity-criteria are obscure; necessary coextensiveness provides our only handle on the notion of sameness of kinds (cf. Mellor [1977]). But we have already seen where that leads us. Teller [1977]

suggests that we take the referents of these terms to be *properties* (determinables and determinates), since property-talk provides convenient paraphrases for talk about kinds of things (animals, substances, colors, shapes, etc.). Inasmuch as Putnam [1970a] defends the need for quantification over properties, let us explore the consequences of saying, on the one hand, that underlying trait terms (*qua* singular terms) refer to certain properties, and, on the other hand, that the extension of such a term at a world is just the extension of the corresponding property at that world.

The idea, then, is that properties of various sorts are the "underlying traits" evoked by underlying trait terms. In introducing, say, 'gold' via the indexical description 'stuff having the same structure as this', one picks out (in a non-identifying way) what we today would describe as the property *P* of being composed of atoms of such-and-such internal structure (this of course depends upon there being such a structure present in the paradigm). The *reference* of 'gold' in future subject-uses is fixed as *P*; the *extension* of 'gold' at a world *w* is the set of all those things existing in *w* which instantiate *P*. Rigidity of reference for mass and count nouns is defined exacly as it was for singular terms like proper names: *N* is a *rigid* designator iff *N* refers to some actually existing item, *X*, and also refers to *X* in every possible world in which *X* exists. (Where, as in the present case, *X* is a property, we may construe existence at a world as having instances in that world; gold doesn't exist in certain possible worlds precisely because nothing instantiates *P* in those worlds.) That 'gold' *is* a rigid designator in this sense follows from the ways its real-world reference is fixed, together with Kripke's counterfactual construal of possible-worlds talk: in modal reasoning about *gold*, we are always talking about *P*, although we allow that *P*'s extension may vary from one counterfactual circumstance to another. That equations like 'Water = H_2O' are necessary if true at all likewise is an immediate consequence of these twin theses about reference-fixing and the nature of modal reasoning. Moreover, all these gains are purchased without paying the price of Essentialism demanded by Cook. To be sure, a modest sort of Essentialism does follow: the necessary truth of 'Water = H_2O' yields the truth of '$\Box\forall x$(x is water $\equiv$ x is H_2O)'--i.e., the claim that being composed of H_2O-molecules is the "essence" of water. But the strong Essentialism about kind-*membership* embodied in Cook's (D-ii) does not follow. That 'gold' rigidly designates *P* does not imply that the members of *P*-kind are necessarily such. (For a brilliant discussion of the failure of attempts to derive Essentialism of this sort from the New Theory of Reference, see Salmon [1979].) Gold may indeed be everywhere identical with *P*-stuff, but without additional essentialist premisses one cannot conclude that anything which is made of gold could

not have been made of anything else!

Having answered the questions about the reference and rigidity of underlying trait terms in this way, we are in effect grasping the second horn of the original dilemma. Carlson [1977] provides a vast amount of linguistic data to support the contention that "bare plurals" in English are best understood as rigid designators (in our sense) of kinds (or kind-determining properties), <u>regardless</u> of whether the pluralized noun is an underlying trait term like 'cats' or an analytically defined term like 'bachelors'; and his data suggest that a similar treatment of mass nouns is also appropriate. If so, then rigidity of reference marks <u>no</u> crucial distinction between underlying trait terms and analytically defined terms, when these are employed as singular terms. And there is no reason why it should. Opinions to the contrary are based on a failure to separate various plausible but independent essentialist claims from the New Theory itself. This failure manifests itself in the idea, alluded to in various expositions of the New Theory (e.g., Schwartz [1979] and [1980]), that underlying trait terms are just <u>natural kind</u> terms. To ensure this coincidence, one will quickly be led to do just what Cook did: one will build Essentialism into the definition of rigid designation in an attempt to guarantee that <u>only</u> natural kinds can be rigidly designated by mass and count nouns; then one will treat rigidly-thus-defined as the characteristic which separates underlying trait terms from others. We have already seen what is wrong with this approach. What distinguishes natural kind terms from other terms is not their rigidity, since even 'bachelors' can be used rigidly (in our sense); the distinction lies not in <u>how</u> the terms designate but in <u>what</u> they designate. Natural kind terms are just those that designate (usually: rigidly designate) <u>natural kinds</u>. And natural kinds (on our chosen view) are just certain special properties. Just <u>which</u> properties they are, is a matter of controversy. Minimally, we might think of a (first-order) property <u>K</u> as being a natural kind relative to a possible world <u>w</u> iff

(i) <u>K</u> is a nontautologous property (i.e., there is some world <u>v</u> possible relative to <u>w</u> such that something in <u>v</u> lacks <u>K</u>);

(ii) <u>K</u> is nomically nonspecific (i.e., if something has <u>K</u> at a world <u>u</u> nomically compatible with <u>w</u>, then there is a world <u>v</u> nomically compatible with <u>w</u> such that the extension of <u>K</u> at <u>v</u> is nonempty and disjoint from the extension of <u>K</u> at <u>u</u>);

(iii) <u>K</u> is nomically realizable (i.e., <u>K</u> is

instantiated at some world nomically compatible with _w_);

(iv) _K_ is nomically essential (i.e., if an item _x_ instantiates _K_ at a world nomically compatible with _w_, then _x_ instantiates _K_ at every world _v_ such that _x_ exists in _v_ and _v_ is nomically compatible with _w_).[6]

The existence of natural kinds in this sense is in no way prejudged by our advocacy of (CT) and the claim that underlying trait terms are rigid designators. If there are natural kinds in this sense, we may designate them rigidly or nonrigidly--underlying trait terms serving the former function and analytically defined terms serving either.

8. INTENSIONS AND INTENTIONS

The anti-Intensionalism inherent in (CT) is, so to speak, practical rather than theoretical. The claim is not that certain words "could not" have intensions; rather the claim is just that (a) otherwise competent ordinary speakers are for the most part unable to give anything faintly resembling a noncircular analytic definition of these terms, and (b) we can nevertheless account for their ability to talk about the real world using these terms without having to posit intensions which are represented in the heads of speakers (or in the heads of those from whom they borrow). Since, to put it crudely, the only reason we could have for believing in _F_s is that we either bump into _F_s or must posit _F_s to explain something connected to what we do bump into, (a) and (b) would seem to provide an excellent basis for practical eschewal of intensions (at least for the words in question).

One could, of course, take this modest anti-Intensionalism one step further and attempt to _demonstrate_ that intensions-in-the-head cannot explain certain data readily explained by (CT). Putnam does just this in his famous "Twin-Earth" argument (Putnam [1975a]). We are to imagine a planet (Twin-Earth) which is exactly similar to Earth at the macro-level but which differs from it in certain unobservable microstructural features. In particular, the liquid which fills the oceans, lakes, and rivers of Twin-Earth--which the Twin-Earthers call 'water'--is, though indistinguishable from H_2O in all observable respects, nonetheless distinct from H_2O, having instead chemical structure XYZ. Putnam claims that (at least in prescientific times) Earthmen's term 'water' has a quite different extension than the Twin-Earthers' term 'water', since each group uses its term to talk about the _local_ liquid, and these liquids are in fact chemically distinct. By hypothesis, (prescientific) Twin-Earthers have exactly the

same beliefs, intentions, etc., regarding their liquid that (prescientific) Earthmen have regarding our liquid. But since these mental states are the material for intensions-in-the-head, it seems to follow that the intensions-in-the-head account of the determination of extension is impotent to explain the evident *difference* of extension, a fact which is readily explained by the supposition that 'water' is an underlying trait term for both Earthmen and Twin-Earthers.

A number of authors (e.g., Linsky [1977], Ware [1978], Zemach [1976]) have taken issue with the Twin-Earth argument on the ground that it presupposes an overly narrow conception of "mental states". Putnam explicitly restricts his discussion to mental states "in the narrow sense", such a state being one which '... presupposes the existence of [no] individual other than the subject to whom that state is ascribed' (Putnam [1975a], p. 220). Now it may be admitted that intensions-in-the-head cannot accomodate the Twin-Earth data when construed solipsistically as mental states in the narrow sense; but, the critics urge, this shows only that a would-be Intensionalist should not construe intensions-in-the-head in this restrictive way! Why not, they suggest, appeal to mental states of a *relational* character? After all, if a speaker uses a term *T* seriously, he presumably intends (*inter alia*) to be using *T* in the same way as his knowledgable contemporaries, to refer to whatever the knowledgable members of his community use *T* to refer to. So, e.g., the typical user of 'lemon' *does* have "in his head" the relational intention to refer to what *the relevant experts in his community* would call 'lemons' (or some similar relational intention regarding his ancestors, teachers, or whomever). But then the Earthman and the Twin-Earther *do* have different mental states after all, since they have beliefs/intentions regarding *different* groups of people (Earth experts *vs*. Twin-Earth experts). This supposedly opens the way to consideration of divergent *speaker-meanings* for various uses of such terms, and--to the extent that linguistic meaning can be defined in Gricean fashion--perhaps even to a full-blown reinstatement of Intensionalism.

Insofar as intensions-in-the-head are capable (at least in principle) of linguistic formulation, we may engage in semantic ascent and clarify what is at issue in this debate without becoming bogged down in controversies over the nature and individuation of mental states. To put it baldly, what the Twin-Earth argument shows is that the totality of *nonindexical* predicates mutually available to prescientific Earthmen and Twin-Earthers is insufficient to *define* 'water' in such wise that 'water'-on-our-definition picks out H_2O while 'water'-on-their-definition picks out XYZ. Putnam never denies that the parties in each world can form *indexical descriptions* which pick out just the substances in question; indeed, it is crucial to his account of 'water' as an

underlying trait term that one and the same indexical description _should_ function on Earth to pick out H_2O and on Twin-Earth to pick out XYZ. But such indexical descriptions, though they can be used to fix the reference of 'water' _at a given index_ (Earth _vs_. Twin-Earth), cannot be profitably employed in an analytical definition of 'water'. For analytical definitions make the definiendum synonymous with the definiens. The extension of the definiendum at any world must be just that of the definiens at that world; but indexicals in the definiens will often shift their reference from world to world, bringing about unwanted changes in the extension of the definiendum. If, e.g., 'water' was simply definitional shorthand for 'whatever the relevant experts at time _t_, place _p_, and world _w_ call 'water'', we would indeed get the desired result that 'water' in our mouths has a different extension from 'water' in Twin-Earthers' mouths, but only at the price of being unable to entertain even the possibility that our experts might be wrong! For at any index (_t_, _p_, _w_), the sentence 'Water is just whatever the current local experts call 'water'' will come out true (or at least not false). The critic's injection of relational attitudes and divergent speaker-meanings comes to no more than the proposal that each ordinary user of 'water' (here or on Twin-Earth) mentally associates with it some such indexical description on each occasion of use, and that this description determines the extension of 'water' as used on that occasion. But this merely relocates the awkward consequence noted above, shifting it from the language as a whole to each speaker's ideolect. Instead of definitional shorthand in English, we have definitional shorthand for a speaker on an occasion; and instead of unwanted analyticities in English we have equally unwanted analyticities-for-a-speaker.

The critic is thus faced by a dilemma. If, on the one hand, 'water' is defined on each occasion of its use by the indexical description associated with it on that occasion, then much plausible counterfactual reasoning using 'water' must become unintelligible to the user. If, on the other hand, the associated indexical description is used merely to fix the reference of 'water' in context and is not held to be interchangeable with it, then we have a view very much like (CT)--except that speakers will be constantly "reintroducing" the term 'water', virtually every successive use amounting to a reintroduction (Goldman [1979] actually proposes such a view). In short, either the critic is just rejecting the Kripke/Putnam intuitions about counterfactual reasoning out of hand and thus securing Intensionalism by brute force, or else his proposal does not constitute an Intensionalist alternative at all. Until some argument is forthcoming to show that the intuitions in question are unsound--and I know of no such argument--Putnam's Twin-Earth argument stands untouched.

Indexical descriptions-in-the-head (the counterpart in our discussion to "relational mental states") may indeed play a role in fixing the extension of terms like 'water'; Putnam's point is that the most plausible construal of how they do so offers no comfort to Intensionalism.[7]

9. THE INTENSIONALIST STRIKES BACK

At least one Intensionalist has tried to turn the table on Putnam by arguing that there are numerous facts about alleged underlying trait terms which can only be explained by invoking intensions. Moravcsik [1981] claims that (CT) and relevantly similar anti-Intensionalist views of certain terms cannot account for (i) the explanatory role of these terms in predications, (ii) their connection with principles of persistence and individuation, (iii) the way these terms are learned, and (iv) the way they function in belief-contexts. Let us consider these complaints in the order listed.

Regarding (i), Moravcsik writes,

> The construal of natural-kind terms as rigid designators seems intuitively plausible when we consider the roles these terms play in subject position. They do, however, play equally important roles in predicative position. Especially significant is their role in "because"-clauses. For example, in using the phrase 'because it is made of iron' to explain the solidity of a structure, we rely on what will be construed as the essential properties of iron (or, in different words, associated criteria) to have our claim interpreted as having explanatory value. Without the link between natural-kind terms and associated properties, it is difficult to assign to these expressions the role they play in typical everyday or scientific explanations. In this respect ... they differ from such paradigms of rigid designators as proper names. (p. 9)

The "link with associated properties" is, of course, supposed to be provided via the intension of the term in question. But this objection simply overlooks Putnam's detailed discussion of the role of Stereotypes in communication (Putnam [1975a], pp. 247-52). To begin with, it is simply false that a connotationless term cannot contribute explanatory value: contributing a referent can have considerable explanatory force in context. Bare demonstratives and demonstrative phrases employing "dummy" nouns like 'thing' or 'stuff' are certainly good candidates for being connotationless designators. But, in a suitable context, I may surely explain the weight,

solidity, etc., of an object (e.g., an enamelled skillet) by saying 'Because it's made of this stuff' and ostending, say, a lump of pure iron. What makes the context "suitable" is just that my audience is antecedently familiar with the ostended stuff. The case of 'Because it's made of iron' is no different in principle: I rely minimally on the audience's grasp of the Stereotype associated with 'iron' (and perhaps additionally on special knowledge they possess--e.g., if I am addressing a group of ironmongers). In my use of 'iron' I do not semantically attribute any properties at all; yet I can communicate to my audience various information about properties because we share a common fund of stereotypical lore about the stuff so-named. Moravcsik seems to confuse what is communicated in an utterance-act with what is semantically entailed by the utterance object (sentence uttered). What needs to be noticed is just that one's success in explaining something to a given audience can hinge as much on the former as it does on the latter.

It must also be remembered that explanation is a highly interest-relative notion. Explanations which employ underlying trait terms and rely on the associated Stereotypes are thus vulnerable from two related directions. On the one hand, the Stereotype may be too impoverished to supply the relevant information. If I attempt to explain why this bow is so supple by saying 'Because it's made of yew', I will probably fail to communicate anything enlightening simply because the Stereotype for 'yew' is well-nigh empty ("a kind of wood"). On the other hand, even a rich Stereotype may not suffice to provide an explanation of the proper "depth". Stereotypical features are usually "superficial" characteristics, but the audience's interests may run much deeper. In pointing out that the skillet is so heavy because it's made of iron (where, let us suppose, "heavy metal" is part of the Stereotype for 'iron'), I am offering a relatively superficial explanation which will be of interest only to someone who (a) knows the Stereotype and (b) did not know that the skillet was made of iron. If, however, the audience knows what the skillet was made of but wants to know why iron objects are heavy, then the stereotypical heaviness of iron is not to the point.

These considerations, however, provide no aid to Moravcsik's criticism. Suppose I have some bits of metal which melt in my hands when I pick them up and quickly resolidify when returned to a cool surface. Now in some sense it is a good explanation of this phenomenon to point out that those bits of metal are composed of Gallium. One who offers such an explanation cannot simply rely on the Stereotype for 'Gallium', since (for nonscientists) it includes little more than the information "soft, bluish-white metal". Even if it included "has a low melting point" it would still be too superficial: we can see that the stuff in question melts at

low temperature--what we want to know is what it is about that stuff in virtue of which it melts so readily. But if 'Because it's made of Gallium' admittedly neither entails nor stereotypically communicates the possession of the structural properties that explain its low melting point, how can it be an explanation? Before rushing to the conclusion that it must after all entail the possession of those properties, we should ask ourselves just <u>who</u> is passing judgment on whether 'Because it's made of Gallium' constitutes an explanation. Moravcsik simply assumes the position of the omniscient Ideal Observer, who of course accepts the remark as explanatory because he can fill in all the intermediate steps. But our ordinary explanatory remarks are addressed to lamentably non-Ideal audiences: we are attempting to explain something <u>to them</u>, in the context of (what we take to be) their background knowledge and interests. In the present case, we would offer 'Because it's made of Gallium' as an "explanation" only if we thought that the audience had sufficent scientific knowledge or innate intelligence that, once the element was identified for them, they could look up, remember, or work out for themselves the connection between the structure of Gallium and its peculiar dispositional properties. In this respect the case is exactly like the one involving bare ostensive identification: the "link with associated properties" derives from the audience's background knowledge. The word 'Gallium' tells them where to look, not what they will find.

Regarding (ii), Moravcsik writes,

> Natural-kind terms can form parts of quantified phrases, as in 'some water was needed to revive him'. ... Such quantifications require that natural kinds have articulation. Such articulation is expressed partly by the principles of individuation and persistence that we attach to them. ... Principles of individuation and persistence, however, are parts of <u>intensions</u>. Knowing associated principles of persistence and individuation is part of what it is to know and understand a language. Furthermore, this aspect of understanding cannot be segregated from matters of reference, since the principles of individuation yield the domain of entities to be referred to. (pp. 10-11)

This objection fails to appreciate that one of the important purposes for which we have underlying trait terms in our language is precisely to enable us to <u>postpone</u> sticky questions about individuation and persistence of kinds while provisionally getting on with the business of talking about these kinds. Consider, e.g., the element gold. The persistence and individuation of an element are presumably matters which hinge

on the identity of its atomic structure: no element is gold unless it has the appropriate underlying structure, and gold persists so long as that structure is exemplified. (Questions about the kind gold must be distinguished from questions about objects made of gold. The individuation and persistence of golden things may be more or less independent of questions of composition: e.g., on some views, the Golden Calf may persist while its composition gradually changes through a series of cut-rate repairs using cheap alloys.) But 'gold' was used to talk about gold long before anyone knew anything about atomic structure, hence long before the appropriate criteria of individuation and persistence for the element itself could be given. The introducer of an underlying trait term like 'gold' leaves it to others (e.g., to future scientists) to identify the underlying trait which "determines" the kind and in terms of which its individuation and persistence are ultimately to be understood. As we have repeatedly stressed, one does not have to know what one is talking about (in any "deep" sense of "knowing what") in order to be talking about it. That 'principles of individuation yield the domain of entities to be referred to' is certainly true, but one need not already have the former to get at the latter, provided that we are willing to admit that the ultimate objects of our discourse might turn out to be very different from our stereotypical conceptions of them. And the genuineness of this possibility is a cornerstone of the New Theory of Reference which nothing so far has dislodged.

Moving on to (iii), we find Moravcsik claiming that

> We do not learn natural-kind terms the way we learn proper names or deictic terms (demonstratives); in order to account for the learning of natural-kind terms we need to posit intensions. The relation between the child's use of 'father' and the use of the same word by an adult is not one of co-reference to the exact same natural kind, but rather one of overlapping criteria. ... [T]here may be situations ... [e.g., where the man is head-of-family but not the biological parent]... in which the child would not know what to say, but which would be accommodated by adult usage in view of adult criteria. (p. 11)

The process of learning natural-kind terms is, we are told, best thought of as one in which the intension initially associated by the child with the term is gradually modified and enriched through supplementary and corrective experiences until it finally approximates or becomes identical with the intension which adults associate with the term. The term 'father' is not the best example Moravcsik could have chosen,

inasmuch as it is genuinely ambiguous as between two construals: one, as an underlying trait term or hybrid paraphrased by 'male (biological) parent'; and the other, as a "one-criterion" word analytically defined by something like 'legal male parent'. Kinship-terms in general are poor examples for the Intensionalist precisely because their usage typically antedates the discovery by the primitive speech-community of causal connections between sexual intercourse and pregnancy. Kinship-terms are thus prime examples of Putnam's thesis that "one-criterion" words tend to acquire "natural kind senses" (i.e., to be reintroduced as underlying trait terms) under the pressure of new information which suggests that "something deeper" is at work. 'Father' most likely started out as a role-word like 'hunter' or 'chief'. The child's linguistic progress invites similar treatment: 'father', as a one-criterion role-word in his vocabulary, gradually grows a second "natural kind" use as he is inculcated into the mysteries of sex. Notice, however, that this growth is not very plausibly viewed as simply adding another layer to the intension of 'father' as previously used. For most adults who use 'father' in the "natural kind" way are woefully ignorant of the biology of reproduction; though perhaps not a pure underlying trait term for them, 'father' is at best a hybrid, defined by something like 'male parent who has produced offspring'--where 'produced offspring' is a pure underlying trait expression in their vocabulary. Of course if the child is using 'father' in the one-criterion way and his parents are using it as an underlying trait terms or hybrid, then they are not referring to the same kind at all, but this shows nothing about the need to posit intensions for natural kind terms. Similar remarks apply, mutatis mutandis, to gestalt-terms as well as to role-words. Probably very few (if any) of our ordinary, nontechnical underlying trait terms originated as such; to the extent that one-criterion words persist or are newly created, talk of intensions has a clear point in explanations of how they are learned. And since underlying trait terms are frequently introduced by reference-fixing descriptions containing some one-criterion words (which may or may not survive in their own right), intensions will also have a role in explaining the teaching and learning of underlying trait terms. But none of this shows that underlying trait terms themselves must be regarded as having intensions which are somehow represented in the heads of those who use them.

Finally, a brief remark about (iv). Moravcsik never explicitly formulates "the problem of belief-contexts"; but, in light of the few hints he gives, it is highly likely that what he has in mind is the following objection. Insofar as a belief-sentence is taken purely de dicto, the embedded clause (dictum) giving the "content" of the belief in question is notoriously not open to substitution of co-referring singular

terms without peril to the truth-value of the whole belief-sentence. We have seen that underlying trait terms can function as singular terms. When they are so used in the _dicta_ of belief-sentences, they too do not tolerate replacement of co-referring members of their kind. For, it is claimed, 'John believes that _water_ is tasty' may be true while 'John believes that _H_2O_ is tasty' is simultaneously false, despite the truth (indeed, the necessary truth) of 'Water = H_2O'. Now if, as the New Theory has it, there is nothing _semantically_ different about 'water' and 'H_2O' (both being connotationless rigid designators of the same substance), then, assuming compositionality, substitution _salva veritate_ ought to be the rule. But it isn't. So there must after all be some semantic difference between 'water' and 'H_2O'. And, the Intensionalist asserts, this significant semantic difference can only be explained as a difference in the _intensions_ of 'water' and 'H_2O'.

This objection would be cause for some alarm were it not for some recent work of Kripke's (Kripke [1979]). The crucial premiss of the objection is the one which asserts that sentences of the forms ⌜_X_ believes that $(\ldots \underline{T}_1 \ldots)$⌝ and ⌜_X_ believes that $(\ldots \underline{T}_2 \ldots)$⌝ may differ in truth value even though $\underline{T}_1$ and $\underline{T}_2$ are rigidly codesignative underlying trait terms. What Kripke shows is that the commonsensical principles of belief-attribution which lead us to accept this premiss are _inconsistent_, hence that we have no good reason to regard it as true. This, of course, does not show the premiss to be false, but it does deprive the objector of any pretension to having a "knock-down proof" of the alleged substitution-failure. Let us quickly recapitulate Kripke's reasoning as it would apply to underlying trait terms.

The crucial ingredient in these belief-attributions---which Kripke calls the "Disquotation Principle"--may be stated (for English) as follows (where '$\Pi \underline{p}$' is a substitutional quantifier whose substitution-class is the set of non-indexical English declaratives):

(DPE) $\Pi \underline{p} \forall \underline{x}$($\underline{x}$ is an English-speaker who sincerely, reflectively, and literally utters or assents to '$\underline{p}$'$\rightarrow \underline{x}$ believes that $\underline{p}$).

Corresponding Disquotation Principles may be obviously stated in and for other languages (with appropriate reconstruals of the initial quantifier). Now the standard reason offered for holding that, e.g., 'John believes that water is tasty' can differ in truth-value from 'John believes that H_2O is tasty' is that John may assent (sincerely, etc.) to 'Water is tasty' and 'H_2O is _not_ tasty', whence by (DPE) it follows that he believes that water is tasty and that H_2O is not tasty, whence--by the Tarski-biconditional for the truth-predicate and the plausible assumption that John believes no _explicit_

contradictions--it follows that 'John believs that water is tasty' is true but 'John believes that H_2O is tasty' is false.

Now consider Hans, a monolingual German speaker. Hans learns '<u>Asbest</u>', the German word for asbestos, in a situation where he is shown a theater fire curtain and told that the name refers to the kind of stuff from which the curtain is made. Being young and naive, Hans supposes that the fibrous stuff in question is some sort of <u>animal</u> <u>hair</u>; and since no one around him knows better, his error never gets corrected. In particular, Hans assents (sincerely, etc.) to 'Der Asbest ist kein Mineral' and even assertively utters it on occasion. It is therefore correct to say (in German)

(1) Hans sagt: 'Der Asbest ist kein Mineral'.

Reasoning in German, we use the German version of the Disquotation Principle to pass from (1) to

(2) Hans glaubt, dass der Asbest kein Mineral ist.

Then, using the German version of the Tarski-biconditional (with German as object-language and extended German as meta-language), we pass from (2) to

(3) 'Hans glaubt, dass der Asbest kein Mineral ist' ist wahr.

Now (3) translates into English as

(4) 'Hans glaubt, dass der Asbest kein Mineral ist' is true.

But the German sentence of which truth is thus predicated itself translates as

(5) Hans believes that asbestos is not a mineral.

So by the principle that correct translation of nonindexical declaratives is truth-value-preserving, we infer

(6) 'Hans believes that asbestos is not a mineral' is true.

And from (6) we then obtain

(7) Hans believes that asbestos is not a mineral.

So far, so good. But now Hans travels to America to learn English by the total immersion method. One day he encounters some unfamiliar looking rocks and asks what they are called. The answer, of course, is 'asbestos'. We may thus think of the reference 'asbestos' being fixed for Hans by the description 'this kind of mineral'. Hans now (sincerely, etc.) assents to and asserts

(8) Asbestos is a mineral.

But from this fact we can use (DPE) to obtain

(9) Hans believes that asbestos is a mineral.

Here we face a nasty paradox. In light of the theory, it seems wildly implausible to claim that either or both of (7) and (9) are false. Nor can we simply bite the bullet and <u>accept</u> both (7) and (9), claiming that poor Hans "just has contradictory beliefs". For although a person certainly can have contradictory beliefs, he could in principle, if he knew enough logic, come eventually to <u>see</u> that they are

contradictory. But no amount of logical acumen on Hans' part will allow him to detect any contradiction between 'Der Asbest ist kein Mineral' and 'Asbestos is a mineral'. What he lacks is not reasoning power but a crucial bit of metalinguistic information: viz., that 'der Asbest' (in German) and 'asbestos' (in English) refer to the same mineral.

The alleged "proof" of substitution-failure thus collapses for want of support, taking down with it the allied argument for attributing intensions to underlying trait terms. This is not to deny that there are still difficulties with the idea that connotationless rigidly codesignative terms are intersubstitutable in belief-contexts (see Boër [forthcoming]).[8] But the substitution-issue now has the status of an open problem for further research. Both the critics and the proponents of the New Theory must await the outcome.

NOTES

1. Intensionalists often try to ameliorate the difficulty by reserving intension-determining status for those beliefs which are "most central" to our picture of the things in question. The problem with this ploy is that it is little more than an attractive metaphor. Centrality is presumably a matter of importance relative to local interests, and importance is a matter of degree. Just how important relative to other beliefs does a given belief have to be in order to be intension-determining? And why draw the line at one point rather than at another nearby? Given the variety and mutability of human interests, the Intensionalist faces the unwelcome task of steering a course between sheer ad hoc stipulation on one side and the anarchy of linguistic solipsism on the other.

2. There is a striking parallel between the behavior of 'water' in such a case and that of certain proper names and definite descriptions, which may not only have psychologically determined "speaker's referents" distinct from their historically determined "semantic referents" but may also undergo a de facto reintroduction which refixes the latter in terms of the former. For detailed discussion, see Evans [1973] and Kripke [1977].

3. The notion of the Stereotype associated with a term has undergone a number of revisions in Putnam's writings (especially as between Putnam [1970b] and Putnam [1975a]). For the purposes of this paper, we shall construe a term's Stereotype in a given speech community to comprise that

community's basic common lore about the things or stuff so denominated. In other words, the community's Stereotype for a term _T_ comprises the community's "stock" answer to the question ⌜What is (a) _T_?⌝, the answer that any minimally competent speaker of the language is expected to be able to produce. Stereotypes may be very superficial and/or inaccurate; for they merely embody what, relative to community-wide interests, is found "striking" about the (possibly atypical) objects of which the community has had experience. Stereotypes are in no way analytically tied to terms. Thus, e.g., 'water' has (for ordinary English speakers) a comparatively rich Stereotype as a clear, colorless fluid found in such-and-such places and useful for certain standard purposes; but it is perfectly conceivable that water should cease to have any or all of these characteristics (e.g., if various natural laws were altered). Putnam sometimes distinguishes the Stereotype for a term from its "Semantic Marker", the latter being a general category-indicator such as "animal", "mineral", "liquid", etc., and the former being regarded as giving various putative _differentia_ within the category. However, Putnam ([1973], p. 205) stresses that Semantic Markers are no more analytically tied to terms than are their Stereotypes: it is not analytic that water is a liquid, that tigers are animals, etc. This textual fact is often overlooked by critics; thus both Goosens [1977] and Mellor [1977] seem to accuse Putnam of holding just the opposite view! The critics have probably been misled by Putnam's loose and deliberately ironical use of the word 'meaning' in his discussion (Putnam [1975a]) of the "canonical form for meaning specifications". For the purposes of this paper, Stereotypes will be regarded as already including such category-indicators as may be appropriate in context.

4. Wiggins [1980] worries unduly about the status of theoretical terms introduced to refer to things whose existence is predicated by a theory, but which have not yet been discovered or created--i.e., which have at the time of introduction no paradigms (cf. esp. pp. 210-213). His problem arises from identifying Putnam's view with (MT) and not seeing that Putnam goes on to espouse a more comprehensive view like (CT), which avoids the difficulty at the outset.

5. In all fairness to the critics, it must be admitted that Putnam often talks in ways which invite this misunderstanding. In particular, his choice of examples of common names of plants and animals heavily favors those which correspond fairly closely to recognized biological taxa.

6. Where natural kinds are concerned, the relevant notions of possibility and necessity would appear to be nomic rather than alethic. (i) and (ii) are meant to rule out "trivial" properties like being-red-or-nonred and being-identical-to-Socrates-or-Plato (cf. Browning [1978]). (iii) and (iv) are adapted from Cocchiarella [1976], where a number of additional qualifications are proposed. By relativizing natural-kindhood to a world, I have left open the question of whether a property K which counts as a natural kind relative to our world (and hence relative to all worlds nomically compatible with ours) also counts as a natural kind relative to some or all worlds not nomically compatible with ours. I suspect the answer is negative, but it would take us too far afield to pursue the answer here.

7. This is not the place to argue the merits of Methodological Solipsism, but it should be noted that if we are willing to grant the correctness of Methodological Solipsism then there is a very powerful objection we can make against the "relational attitudes" ploy. For insofar as these relational attitudes can differ while the mental representations functioning within the subject (by reference to which his behavior is to be explained) remain the same, the posited relational attitudes will be of no explanatory value to psychology narrowly construed.

8. In particular, there is the problem of explaining why, if English permits such substitutions, it is not readily apparent to any semantically competent English-speaker that it does! The cited paper addresses this problem at length.

REFERENCES

Boër, S.: forthcoming, 'Names and Attitudes', in M. Salmon (ed.), _The Philosophy of Logical Mechanism_, D. Reidel, Dordrecht, Holland.

Browning, D.: 1978, 'Believing in Natural Kinds', _The Southwestern Journal of Philosophy_ 9, 135-148.

Carlson, G.: 1977, _Reference to Kinds in English_, University of Massachusetts Doctoral Dissertation, reproduced by Indiana University Linguistics Club.

Cocchiarella, N.: 1976, 'On the Logic of Natural Kinds', _Philosophy of Science_ 43, 202-222.

Cook, M.: 1980, 'If "Cat" is a Rigid Designator, What Does it Designate?', _Philosophical Studies_ 37, 61-64.

Dampier, W.: 1929, _A History of Science_, Cambridge University Press, Cambridge.

Donellan, K.: 1973, 'Substances as Individuals', _Journal of Philosophy_ 70, 711-712.

Dupré, J.: 1980, 'Natural Kinds and Biological Taxa', _The Philosophical Review_ 90, 66-90.

Enc, B.: 1976, 'Reference of Theoretical Terms', _NOUS_ 10, 261-282.

Evans, G.: 1973, 'The Causal Theory of Names', in Schwartz [1977].

Fine, A.: 1975, 'How to Compare Theories: Reference and Change', _NOUS_ 9, 17-32.

Goldman, A.: 1979, 'Reference and Linguistic Authority,' _The Southern Journal of Philosophy_ 17, 305-321.

Goosens, W.: 1977, 'Underlying Trait Terms', in Schwartz [1977].

Hollinger, R.: 1974, 'Natural Kinds, Family Resemblances, and Conceptual Change', _The Personalist_ 55, 323-333.

Kornblith, H.: 1980, 'Referring to Artifacts', _The Philosophical Review_ 89, 109-114.

Kripke, S.: 1972, 'Naming and Necessity', in D. Davidson and

G. Harman (eds.) Semantics of Natural Language, D. Reidel, Dordrecht, Holland.

Kripke, S.: 1977, 'Speaker's Reference and Semantic Reference', Midwest Studies in Philosophy 2, 255-276.

Kripke, S.: 1979, 'A Puzzle About Belief', in A. Margalit (ed.) Meaning and Use, D. Reidel, Dordrecht, Holland.

Linsky, B.: 1977, 'Putnam on the Meaning of Natural Kind Terms', Canadian Journal of Philosophy 7, 819-828.

Mackie, J.: 1974, 'Locke's Anticipation of Kripke', Analysis 34, 177-180.

McGinn, C.: 1975, 'A Note on the Essence of Natural Kinds', Analysis 35, 177-183.

McGinn, C.: 1976, 'On the Necessity of Origin', Journal of Philosophy 73, 127-135.

McKay, T. and C. Stern: 1979, 'Natural Kind Terms and Standards of Membership', Linguistics and Philosophy 3, 27-34.

Mellor, D.: 1977, 'Natural Kinds', British Journal for the Philosophy of Science 28, 299-312.

Moravcsik, J.: 1981, 'How Do Words Get Their Meanings?', Journal of Philosophy 78, 5-24.

Pratt, V.: 1972, 'Biological Classification', British Journal for the Philosophy of Science 23, 305-327.

Putnam, H.: 1962, 'The Analytic and the Synthetic', in H. Feigl and G. Maxwell (eds.) Minnesota Studies in the Philosophy of Science, vol. 3, University of Minnesota Press, Minneapolis (reprinted as chapter 2 of Putnam [1975d]).

Putnam, H.: 1970a, 'On Properties', in N. Rescher et al. (eds.) Essays in Honor of Carl G. Hempel, D. Reidel, Dordrecht, Holland (reprinted as chapter 19 of Putnam [1975c]).

Putnam, H.: 1970b, 'Is Semantics Possible?', in H. Kiefer and M. Munitz (eds.) Language, Belief, and Metaphysics, State University of New York Press (reprinted as chapter 8 of Putnam [1975d]).

Putnam, H.: 1973: 'Explanation and Reference', in G. Pearce and P. Maynard (eds.), Conceptual Change, D. Reidel, Dordrecht, Holland (reprinted as chapter 11 of Putnam [1975d]).

Putnam, H.: 1975a, 'The Meaning of "Meaning"', in K. Gunderson (ed.) Language, Mind and Knowledge, University of Minnesota Press, Minneapolis (reprinted as chapter 12 of Putnam [1975d]).

Putnam, H.: 1975b, 'Language and Reality', in Putnam [1975d], chapter 13.

Putnam, H.: 1975c, Mathematics, Matter, and Method: Philosophical Papers, vol. 1, Cambridge University Press, Cambridge.

Putnam, H.: 1975d, Mind, Language and Reality: Philosophical Papers, vol. 2, Cambridge University Press, Cambridge.

Putnam, H.: 1978, Meaning and the Moral Sciences, Routledge and Kegan Paul, Boston.

Salmon, N.: 1979, 'How Not to Derive Essentialism from the Theory of Reference', Journal of Philosophy 76, 703-725.

Schwartz, S. (ed.): 1977, Naming, Necessity and Natural Kinds, Cornell University Press, Ithaca, N.Y.

Schwartz, S.: 1978, 'Putnam on Artifacts', Philosophical Review 87, 566-574.

Schwartz, S.: 1979, 'Natural Kind Terms', Cognition 7, 301-315.

Schwartz, S.: 1980a, 'Natural Kinds and Nominal Kinds', Mind 89, 182-195.

Schwartz, S.: 1980b, 'Formal Semantics and Natural Kind Terms', Philosophical Studies 38, 189-198.

Teller, P.: 1977, 'Indicative Introduction', Philosophical Studies 31, 173-195.

Ware, R.: 1978, 'The Division of Linguistic Labor and Speaker Competence', Philosophical Studies 34, 37-61.

Wiggins, D.: 1980, Sameness and Substance, Cambridge University Press, Cambridge.

Zemach, E.: 1976, 'Putnam's Theory on the Reference of Substance Terms', Journal of Philosophy 73, 116-127.

Frank Jackson

THE EASY EXAMINATION PARADOX

Surprise examinations in a specified period are possible, moreover it need be no surprise that they are a surprise. That much almost everyone has found obvious. Call it 'the strong intuition'. In this paper I defend the strong intuition against an only too familiar argument discussed in an only too familiar literature. This defence requires that I solve a version of the examination paradox. But we will see that it is, unfortunately, only an easy version of that paradox which I solve. My solution will draw on points made by W.V. Quine,[1] and by Crispin Wright and Aidan Sudbury.[2] I think that Wright and Sudbury got the ingredients of the solution to the easy paradox pretty well right, but not the recipe.

The paper divides up as follows. In section 1 I state a version of the examination paradox. In section 2 I give the ingredients for our solution. In section 3 I say why I am unhappy with Wright and Sudbury's recipe. In section 4 I give my recipe. In section 5 I explain why we have only solved an easy version of the paradox, though we have done enough to vindicate the strong intuition; and conclude by drawing a couple of morals concerning the hard examination paradox.

1. THE EXAMINATION PARADOX

On Monday morning the headmistress of a girls' school announces two things, which I will call the existence condition (EC) and the ignorance condition (IC).

(EC) An examination will be held one afternoon this week (Monday to Friday).

(IC) You won't know until after noon on the day which afternoon it is.

The girls reason that she cannot be speaking truly as follows. "Suppose the examination is held Friday afternoon. Then at Friday noon we would know that Monday through Thursday have proved exam-less and so that, as Friday afternoon is the only one remaining, it must be the one--we would know Friday noon that the exam was Friday afternoon, which violates (IC). Now suppose the exam is held on Thursday afternoon. Then at Thursday noon we will know that Monday through Wednesday have

B. K. Matilal and J. L. Shaw (eds.), Analytical Philosophy in Comparative Perspective, 151–159.

proved exam-less, and so that the exam must be on Thursday or Friday afternoon by (EC). But Friday afternoon is out by the result just obtained, hence Thursday afternoon must be the one. Therefore we would know Thursday noon that the exam was that afternoon, which violates (IC). So Thursday afternoon is out also--and so on for all the afternoons.

The girls' reasoning is fallacious. To see why we need to have three points before us.

2. THE INGREDIENTS

I will call the three points 'Quine' (Q), 'Son of Quine' (S of Q) and 'Defeasibility' (D). (Q) is emphasized in Quine, op. cit.; (S of Q) is emphasized in Wright and Sudbury, op. cit., and arguably is implicit in Quine, op. cit. (D) I have extracted from Wright and Sudbury, op. cit.; I leave it to the reader to assign degrees of responsibility for it.

> (Q) The exam can be held on Friday afternoon without violating (IC) if at Friday noon the girls are in doubt about (EC).

Suppose noon of the last day comes without an exam making its appearance and, perhaps as a result of this fact or perhaps for some other reason, the girls at that time doubt that the exam will be held in the specified period, then the exam can be held on that last afternoon without the girls knowing beforehand. They won't know beforehand because they are not sure that the exam will be held on one of the days specified, and they need to be sure of that to know beforehand.

(S of Q) relates to doubt about (IC) rather than (EC).

> (S of Q) The exam can be held on Thursday afternoon without violating (IC) if the girls are in doubt of (IC) itself Thursday noon.

Suppose that noon of the second last day, Thursday, arrives without the exam; and suppose that, for whatever reason, there is no question of doubting (EC), but that the girls at Thursday noon do doubt (IC). Then they must allow as a possibility that the exam is held on Friday afternoon. True, as (EC) is not to be doubted, doing that would violate (IC), but maybe exactly that will happen. But then the headmistress can hold the exam on Thursday afternoon without violating (IC), for the girls at Thursday noon would still have Friday down as a possibility.

(Q) turns on the fact that doubt about (EC) at Friday noon enables a Friday afternoon exam to satisfy (IC); (S of Q) turns on the fact that doubt at Thursday noon about (IC)--even

in the absence of doubt or potential doubt about (EC)--enables a Thursday exam to satisfy (IC).

(D) concerns a general feature of justified belief.

> (D) One may be justified at t_1 in believing p while acknowledging at t_1 that should certain things happen between t_1 and t_2, one would not be justified at t_2 in believing p.

(D) is not controversial. Here is a simple illustration of it. I am to toss a coin five times. I know that it is a fair coin and that the results of the tosses are probabilistically independent. I am at the start justified in believing that at least one toss will land heads, the probability of this being about 0.97, while acknowledging that should the first four tosses land tails I would not _then_ be justified in believing that at at least one toss will land heads, the probability _then_ being only 0.5.

3. THE WRIGHT-SUDBURY RECIPE

How can we put (Q), (S of Q), and (D) together to solve the (easy) examination paradox? It is immediately clear from (Q) and (S of Q) that the girls need to assume that they are entitled to believe (EC) and (IC) throughout, not just at the beginning. I will call these two assumptions 'belief in EC' (BEC) and 'belief in IC' (BIC).

> (BEC) At every noon prior to the exam, it is reasonable to believe EC.

> (BIC) At every noon prior to the exam, it is reasonable to believe IC.

(Obviously, for both (BEC) and (BIC), if it is reasonable to so believe prior to the exam, it is reasonable to so believe after the exam on the basis of memory and hence to so believe throughout.)

The fact that (BEC) and (BIC) are needed is in itself no solution to the paradox. It is part of the strong intuition that one might be entitled to believe throughout in both (EC) and (IC), as well as that they might both be in fact true. It would be a Pyrrhic victory to answer the girls' reasoning at the cost of making either (BEC) or (BIC) false.

With (BEC) and (BIC) before us, we are in a position to consider Wright and Sudbury's recipe. Here is the crucial passage (changed to conform with our notation and example).

> Suppose that the pupils have neither better reason to believe [EC] than to believe [IC], nor conversely .... In that case they are not entitled to the simplest of the tacit assumptions on which their reasoning depends.... If Friday [noon] comes and no examination has taken place, they will not be entitled reasonably to believe both [EC] and [IC]. Hence the reasonable course, on the present supposition, will be to suspend belief in each.[3]

Wright and Sudbury then go on to consider the other two possibilities--that the evidence for (EC) is better than that for (IC), and conversely--and argue that in those two cases also the paradox can be solved. But I'll start by considering what they say about the case where the evidence for both is on a par.

The prime intuition which we adverted to at the very beginning was that it need not be a surprise that an exam is held in a specified period and is a surprise. More exactly, it is that it is possible that as well as (EC) and (IC) being satisfied, the girls may throughout believe and be justified in believing (EC) and (IC). (Also we may suppose that they justifiably believe that they justifiably believe (EC) and (IC), and also their logical consequences. We don't want to solve the paradox at the cost of making the girls ignorant of themselves or of logic.) Now what Wright and Sudbury claim in the quoted passage is that if the girls' reasons for (EC) and (IC) are on a par, then if noon of the last day arrives without an exam appearing they will have to abandon both (EC) and (IC). Maybe, but that is no solution; it is rather a way of saying that in the case under consideration the last afternoon is not a possible one for the exam. Hold the exam on Friday afternoon and both of (BEC) and (BIC) will be false. That is what we are being told. From which we can infer that the kind of surprise exam intuition says can be held cannot be held on the last day. We have given the paradox-mongers their first step, rather than a refutation!

A similar point applies to their discussion of the case where the evidence for (EC) is better than that for (IC).

> 'But what' it will be protested, 'if the pupils have better reason to believe [EC] than to believe [IC]?' (Suppose, e.g., that they previously checked that the pack contains the Ace of Spades.) In that case, consider their situation on Thursday [noon] ... if the 'examination' has still not occurred. A [Friday] examination would violate [IC]; so if [IC] is true, so is

> [the exam is on Thursday afternoon]. Can it in these circumstances still be reasonable for them to believe [IC]? Supposing that their reason for believing [EC] would in fact survive the non-occurrance of the examination on [Thurday afternoon] ... we have to consider two possibilities respectively: that they have, and that they have not, reason to believe on Thursday [noon] that this will be so.
>
> Suppose, first, that they have reason to believe that [Friday noon] they will still have reason to believe [EC], come the examination [Thursday afternoon] or not. In that case, they can no longer reasonably believe [IC]; ... But now, <u>lacking</u> any reason to believe [IC]; they have no reason to rule out a [Friday afternoon] examination. Hence the way is open for the [headmistress] to fulfill the announcement by setting the examination on [Thursday afternoon].[4]

Yes, the way is open, but at the cost of falsifying (BIC)! The paradox mongers, armed with (BEC) and (BIC), could agree entirely. "You've shown how a Thursday afternoon exam can be held in conformity with (EC) and (IC), but at the cost of it not being rational for the girls to believe (IC) at Thursday noon. And that's a prior noon, so (BIC) is false; so you haven't shown the strong intuition is secure from the girls' reasoning.

4. OUR RECIPE

Once we have added (BEC) and (BIC) to the specification of the paradox, the last day is really out. The exam cannot be held on Friday afternoon. But it can be held on an earlier afternoon consistently with all that has been given. From (D) we have it that the following situation at a noon earlier than Friday is consistent. (i) No exam has yet made its appearance, (ii) the girls are entitled to believe that the exam will occur at some future afternoon, and (iii) the girls acknowledge that should Friday noon arrive without the exam having appeared, they would no longer be entitled to believe that the exam would occur within the specified period. It is important to see that <u>Modus</u> <u>Tollens</u> is beside the point here. The girls would be arguing fallaciously if they argued at, say, Wednesday noon, "We are entitled to (EC); if Friday noon should arrive without an exam, we would not; therefore Friday noon won't." The first premise must be 'We are entitled to (EC) <u>at</u> <u>Wednesday</u> <u>noon</u> (or, <u>now</u>)', while the conditional premise must be 'If Friday noon should arrive without an exam, we

would not be entitled to (EC) <u>at Friday noon</u> (or, <u>then</u>)'. As we remarked before, (D) is not controversial.

Suppose then that the situation described in the preceding paragraph obtains at, again say, Wednesday noon, and suppose that the headmistress puts the exam on that very afternoon. It is clear that all the conditions can be satisfied. The exam is held in the specified period, so (EC) is true. The girls are entitled by (ii) above to believe (EC) at Wednesday noon, and obviously might have been on both Monday and Tuesday. But Monday, Tuesday and Wednesday are all the prior noons, so (BEC) is true. Moreover at Wednesday noon the girls are unable to argue that Friday afternoon is out. For they cannot argue that were Friday noon to come without the exam showing, (IC) would be violated, because they grant that <u>then</u> they would not be entitled to (EC). In short, the combination of (iii) above and (Q) shows that the girls have lost the first step of their <u>reductio</u>. Moreover, they might (and normally would) be aware of this very fact, so that at Wednesday noon they justifiably believe (IC). The same will be true of Monday and Tuesday noons. But they are all the prior noons, so (BIC) is true. So also and obviously is (IC); hence we have shown that (EC), (IC), (BEC), and (BIC) can all be true together. In sum, reflection on (D) shows that even adding (BEC) and (BIC) to one's characterization of the examination paradox does not stop (Q) blocking the girls' <u>reductio</u>.

The argument just given rested crucially on (D), the point that one may be entitled to believe p while acknowledging that were such-and-such to happen one would not then be entitled to believe p, as applied to the girls' justified belief in (EC). A natural question to ask now is whether there is a notion akin to that of reasonable or justified or legitimate belief, but stronger, for which a relevant correlate of (D) fails. Certainty may be such a notion. Suppose we take it as part of the logic of certainty that when it is certain that p at a time and we face the question of what to say should such-and-such turn out to be the case, there are just two possibilities: (a) were such-and-such to turn out to be the case it would still be certain that p, or (b) it is certain that such-and-such will not turn out to be the case.

Now consider the impact of replacing (BEC) by:

> (CEC) At every prior noon, it is certain that (EC).

If we again illustrate with the case of Wednesday noon arriving without the exam showing, the girls now can rule out Friday afternoon <u>at Wednesday noon</u>. For when we put, or they consider, the possibility that Friday noon should arrive without the exam showing, they can say <u>either</u> that that certainly will not happen, in which case they know the exam is before

Friday afternoon, or that should it happen (EC) would still be certain and so (IC) would be violated because (Q) is no threat. In either case, they can say that Friday afternoon is out and so have the first step of their *reductio*.

Nevertheless, even with (CEC) substituted for (BEC) their reasoning is still fallacious, but the mistake occurs at the second step instead of the first. Previously the point turned on (D), (BEC), and (Q); this time it turns on (D), (BIC), and (S of Q).

Consider, say, Tuesday noon, and the following situation:

(i) No exam has yet made its appearance,

(ii) the girls acknowledge that should Thursday noon arrive without the exam, they would no longer be entitled to believe (IC), and

(iii) the girls are certain of (EC).

And suppose the headmistress puts the exam on that very afternoon, Tuesday.

Obviously the girls are or may be certain of (EC) at Monday noon as well as Tuesday. So they are certain at all prior noons and (CEC) is true, along with (EC); and so also is (BIC). For the girls' *reductio* is blocked at its second step. They cannot argue at Tuesday noon that the exam cannot be held on Thursday afternoon. For they cannot argue from the point that were Thursday noon to come around without the exam showing, (IC) would be violated because they grant that *then* they would not be entitled to (IC). In short the combination of (ii) and (S of Q) blocks the second step of their argument. Moreover, they might (and normally would) be aware of this very fact, so that at Tuesday noon they justifiably believe (IC). The same will be true of Monday noon. They are the two prior noons, so (BIC) is true. Therefore we have shown that (EC), (IC), (CEC), and (BIC) may all be true together.

The crucial difference between our recipe and Wright and Sudbury's is that they focus on the position the girls would be in near the end of the week should no exam show by then, while we focus on the position they are in earlier in the week concerning what they at that time say about the position they would be in later in the week should no exam show by then. What would be the case towards the end of the week is in itself irrelevant to our solution; what matters immediately is what the pupils acknowledge earlier about what would be the case later should there be no exam by then. It is that which prevents them imaginatively projecting themselves, so to speak, to an exam-less end of the week and arguing back using (IC).

Earlier we showed how, in view of (D) and (Q), (EC) and (IC) might both be true, and reasonably believed by the girls at every prior noon and so throughout; now we have shown how, in view of (D) and (S of Q), (EC) and (IC) might both be true, (EC) a matter of certainty, and (IC) reasonably believed throughout by the girls. I take this to complete the vindication of commonsense intuitions concerning the examination paradox.

5. THE HARD PARADOX

We have said nothing about what happens if (BIC) is replaced by:

(CIC) At every prior noon, (IC) is certain.

What I take to be the hard examination paradox is got by giving the girls not (BEC) and (BIC), but (CEC) and (CIC). And given the pivotal role of (D) in our discussion together with the plausible claim that no relevant correlate of (D) is true for certainty, I think that it must be concluded that we have not solved the hard paradox; and a similar conclusion goes for Wright and Sudbury even if my criticism in section 3 fails.

Moreover I fear it must also be concluded that we have not gone to the heart of what has worried the many who have worried about the examination paradox. This is not because they have made much of it being certainty rather than reasonable belief in (EC) and (IC) in their characterization of the paradox. Indeed the bearing of the distinction on the paradox is often neglected in the literature. (It is to Wright and Sudbury's credit that their paper makes one aware of its relevance.) The reason rather is that the sort of things we have been saying, and that Wright and Sudbury say, can equally be said about a case no-one (well hardly anyone) has thought to present a particular problem. Here is the case.

Instead of saying one afternoon this week, the headmistress says one afternoon in the next one thousand days. And suppose for whatever reason the girls are certain of this. Instead of saying that the girls will not know until after noon on the day, the headmistress says that no <u>perceptual</u> sign will be given until after noon on the day. This situation, everyone agrees, does not give rise to a paradox. Yet (EC) and (CEC) are true, and it is very highly probable that (IC) and (BIC) are true (.999 and about .998, respectively). The reason there is no puzzle is, precisely, (D). The girls cannot work back from the end of the one thousand days because, although (BIC) is almost certain, it is not quite certain; and consequently should enough days elapse without the exam showing--as just might happen--they would then no longer be entitled to believe (IC).

Nevertheless, what we have said suggests two things about the solution to the hard paradox. First, the trouble is caused essentially by (CIC), and it is noteworthy that (CIC) is not as straightforward as (CEC). There are clear cases in the literature where something like (CEC) is true.[5] This is indeed the standard criticism of Quine's solution. But I know of no clear cases where (CIC) is true. Secondly, as we said before, solving the easy examination paradox is enough to vindicate the strong intuition that everyone feels concerning the possibility of surprise examinations. This means we need not be worried if solving the hard paradox involves recondite matters and sophisticated logical machinery. The reason so many have sought a relatively straightforward, "naive" solution is that they have felt it to be so obvious that the girls' reasoning was fallacious. "How can it possibly be necessary to delve into difficulties of an order with those raised by the classical paradoxes when it is so obvious that it is all a piece of sophistry?" The obviousness attaches to the easy paradox, not the hard one.

Monash University

NOTES

1. W.V. Quine: 1953, 'On a So-Called Paradox', Mind 62, 65-67.

2. Wright, Crispin and Sudbury, Aidan: 1977, 'The Paradox of the Unexpected Examination', Australasian Journal of Philosophy 55, no. 1 (May), 41-58.

3. Op. cit., p. 54.

4. Op. cit., p. 54.

5. See, e.g., Ayer, A.J.: 1973, 'On a Supposed Antinomy', Mind 82.

Krister Segerberg

MODELS FOR ACTIONS

1. INTENSION AND CHANGE

In action, world and agent come together: agent without world is just intention and world without agent just change. To understand action both are needed, agent and world, intention and change.

Let us begin with the world. The world is always in some state or other, the actual state; there are in principle only so many states in which it can be. The actual state changes from time to time, and this change is very much the concern of the agent. Before his action, the world is in a state which we may call the prior state. After his action, and usually to some extent because of it, the world is in a state which we may call the posterior state. The posterior state is almost always distinct from the prior state, but it is not excluded that they are the same; and in some cases they will be. It would be a natural idea to identify the agent's action with the ordered pair of these two states, the prior one and the posterior one. At least this pair may be thought to represent the change the world has undergone in connection with the agent's action. Suggestions of this kind have been made in the literature. For example, it may be held that G.H. von Wright's analysis in [8] is of this general kind. But if such an analysis does something for action, it does nothing for intention. A man who intentionally opens the window may or may not intend to lower the temperature in the room, but in either case a lower temperature may result. The concepts dealt with so far would not seem to allow such distinctions to be made.

Let us now look at the agent. We shall assume that when the agent acts intentionally, then he acts on an intention set (a set of intentions). Moreover, we shall assume that an intention is an intention to bring about an event. How intentions form or agents arrive at intention sets, we do not know. The picture drawn in [7] was roughly this. The agent has one or several intentions: there are certain events he intends to (that is, will try to) bring about. In order to act he must have an intention which is operational in the sense that it can be carried out directly. In the language of [7], there is a routine associated with such an intention, and what the agent does when he carries out this intention is to run that

B. K. Matilal and J. L. Shaw (eds.), Analytical Philosophy in Comparative Perspective, 161–171.

routine; this may or may not *realize* his intention. Taking what we call a *deliberation walk*, the agent will typically add more and more specific intentions until he arrives at one that is operational. If it also has the further property--let us call it the *infimum property*--that realizing that intention would realize all the intentions in the intention set, then he might as well end his deliberation walk and run the routine associated with this intention. In this case we say that the last intention has become *operative*, meaning that this is the intention on which his action is based.

Missing from the account given in the second paragraph of the paper is intention. But in fact the agent's intentions will usually be important to the determination of the posterior state. If the agent is omnipotent, he alone determines it. If he is completely powerless, the world alone determines it. But usually the agent is neither, and then the posterior state is determined by agent and world jointly. A view advanced in [6] is that one may think of the agent as excluding all but a certain set of possible states, for this is in effect what the routine associated with his operative intention does: it limits the choice of posterior states to one of a certain set of states. Which one in fact becomes the posterior state is then determined by the world, that is, nature and, perhaps, other agents.

The philosophy of action thus presented may seem primitive. It has one advantage, however: it can be given a precise explication within the broad tradition of intensional logic. To the task of doing so, we now turn.

2. FORMAL SEMANTICS

By an *intention in* a set U we mean any function from U to U. An *intentional frame* is an ordered pair $= \langle U,C \rangle$, where U is a set and C is a function assigning to each set S of intentions in U and element $w \in U$ an element $C(S,w) \in U$. The elements of U are called (*possible total*) states; U the *domain* or *state space* of ; and C the *change function*. The set of all intentions in U is written $Int\ U$.

The connection between the concept of 'intentional frame' and the informal preamble in Section 1 should be obvious. Define a *proposition in* U as a subset of U, and an *event in* U as a function from U to U. A full analysis of the intuitive concept of 'intention' would presumably show it to be more complicated than that of 'event', yet in our formal modelling we have cast intentions as events. The justification (or excuse) for doing so is our convention to view an intention as an intention to bring about an event.

Next we propose an object language to fit this semantics. Such a language is not uniquely determined by the modelling, but the following one seems natural. The two main categories

of symbols are terms and formulas (the terms to render events/intentions, the formulas to render propositions). Non-overlapping supplies of atomic terms and atomic formulas are assumed. We will use **P** for atomic formulas; **A**, **B** for any formulas; π for atomic terms; α, β for any terms.

Terms can be combined to form complex terms: $\alpha \frown \beta$ ('α and β'), $\alpha \smile \beta$ ('α or β'), and $\bar{\alpha}$ ('not α'). Similarly, formulas can be combined to form complex formulas: $\mathbf{A} \wedge \mathbf{B}$ ('**A** and **B**'), $\mathbf{A} \vee \mathbf{B}$ ('**A** or **B**'), $\neg \mathbf{A}$ ('not **A**'), $\mathbf{A} \supset \mathbf{B}$ ('**A** only if **B**'), $\mathbf{A} \equiv \mathbf{B}$ ('**A** if and only if **B**'). Beyond this we also introduce two event constants, *1* ('the certain event') and *0* ('the impossible event'), two propositional constants, $\top$ ('truth', 'the certain proposition') and $\bot$ ('falsity', 'the impossible propsition'), and, finally five special formula making operators. First among the latter is event identity, =, which is applied to pairs of terms: $\alpha = \beta$ ('α is β' or 'α equals β'). Next we have the intention operator, *Int*, and the realization operator, *Real*, which apply to terms: *Int* α ('α is intended [by the agent]' or 'the agent intends α') and *Real* α ('α is realized'). Lastly, there are two unary modal operators, $\Box$ for universal necessity and $\boxdot$ for pragmatic necessity: $\Box\mathbf{A}$ ('it is universally necessary that **A**') and $\boxdot\mathbf{A}$ ('it is part of the situation that **A**').

In order to interpret this language over an intentional frame $\mathfrak{J} = \langle \underline{U}, \underline{C} \rangle$ we need a valuation in $\underline{U}$: a function $\underline{V}$ which assigns to each atomic term π an event $\underline{V}(\pi)$ in $\underline{U}$, and to each atomic formula **P** a proposition $\underline{V}(\mathbf{P})$ in $\underline{U}$. As in modal logic, we call the triple $\langle \underline{U}, \underline{C}, \underline{V} \rangle$ an intentional model (on $\langle \underline{U}, \underline{C} \rangle$. Suppose that $\mathfrak{M} = \langle \underline{U}, \underline{C}, \underline{V} \rangle$ is an intentional model. Then we define, for each term α, an associated event $\|\alpha\|^{\mathfrak{M}}$ in $\underline{U}$ as follows (dropping the superscript as often as possible):

$\|\pi\| = \underline{V}(\pi)$,

$\|0\|$ = the function $\underline{f}$ such that, for all $\underline{w} \in \underline{U}$, $\underline{f}(\underline{w}) = \emptyset$,

$\|1\|$ = the function $\underline{f}$ such that, for all $\underline{w} \in \underline{U}$, $\underline{f}(\underline{w}) = \underline{U}$,

$\|\alpha \frown \beta\|$ = the function $\underline{f}$ such that, for all $\underline{w} \in \underline{U}$, $\underline{f}(\underline{w}) = \|\alpha\|(\underline{w}) \cap \|\beta\|(\underline{w})$,

$\|\alpha \smile \beta\|$ = the function $\underline{f}$ such that, for all $\underline{w} \in \underline{U}$, $\underline{f}(\underline{w}) = \|\alpha\|(\underline{w}) \cup \|\beta\|(\underline{w})$,

$\|\bar{\alpha}\|$ = the function $\underline{f}$ such that, for all

$\underline{w} \in \underline{U}$,
$\underline{f}(\underline{w}) = \underline{U} - \|a\|(\underline{w})$.

Similarly, we should like to define, for each formula, an associated proposition in $\underline{U}$. However, given the idea behind the new operators *Int, Real,* and ⊡, this is not possible in a straight-forward way. As explained in the introduction, it is vital to distinguish between the world and the agent. The intentional model is meant to model the world, and even though in one sense the agent is in the world, in another sense he is not. Because of the latter fact, the agent does not belong in the model, yet he is needed for the truth definition. Evidently, what the agent supplies is a set of intentions. In ordinary modal logic, formulas have truth-values only relative to indices (possible worlds, corresponding to our states). Here we have to go one step further and relativize to intention sets as well as to state; since some formulas require this, the definition does it uniformly for all. Suppose therefore, with $\mathfrak{M}$ as before, that $\underline{S}$ is any intention and $\underline{w}$ any state. Then we define, for each formula **A**, what it means

for **A** to be <u>true in</u> $\mathfrak{M}$ <u>at</u> $\underline{w}$, given $\underline{S}$, in symbols $\underline{S} \models_w^{\mathfrak{M}} \mathbf{A}$ (as with terms we will drop the superscript as often as possible):

$\underline{S} \models_w \mathbf{P}$ iff $\underline{w} \in \underline{V}(\mathbf{P})$,

$\underline{S} \models_w \top$, always,

$\underline{S} \models_w \bot$, never,

$\underline{S} \models_w \mathbf{A} \wedge \mathbf{B}$ iff $\underline{S} \models_w \mathbf{A}$, and $\underline{S} \models_w \mathbf{B}$,
etc.

$\underline{S} \models_w a = \beta$ iff $\|a\|(\underline{w}) = \|\beta\|(\underline{w})$,

$\underline{S} \models_w$ *Int* a iff $\|a\| \in \underline{S}$,

$\underline{S} \models_w$ *Real* a iff $\underline{C}(\underline{S},\underline{w}) \in \|a\|(\underline{w})$,

$\underline{S} \models_w \Box\mathbf{A}$ iff, for all $\underline{x} \in \underline{U}$, $\underline{S} \models_x \mathbf{A}$

$\underline{S} \models_w$ ⊡**A** iff, for all $\underline{T} \subseteq \underline{\text{Int}}\ \underline{U}$, $\underline{T} \models_w \mathbf{A}$

A formula is <u>valid</u> <u>in</u> an intentional frame if it is true at all states in all models on it; <u>valid</u> if valid in all intentional frames.

It might be argued that, in the last clause of the truth definition, <u>Int</u> $\underline{U}$ ought to be replaced by some suitable subset thereof; this would reflect the possibility that not all possible intentions or intention sets are open to the agent. In

view of Section 1 there are two general restrictions that seem plausible, (i) that $\underline{S}$ be finite, and (ii) that $\bigcap \underline{S} \in \underline{S}$. Let us use the terms 'true in the restricted sense' and 'valid in the restricted sense' instead of 'true' and 'valid' if these restrictions are adopted.

3. LOGIC

The two truth definitions in Section 2 define certain classes of formulas. Let $\underline{L}$ and $\underline{L}^*$ be these classes; that is, the sets of formulas that are valid, respectively, valid in the restricted sense.

Each of $\underline{L}$ and $\underline{L}^*$ is closed under Modus Ponens (if it contains $\mathbf{A} \supset \mathbf{B}$ and $\mathbf{A}$, then it contains $\mathbf{B}$), Universal Necessitation (if it contains $\mathbf{A}$, then it contains $\Box\mathbf{A}$), Pragmatic Necessitation (if it contains $\mathbf{A}$, then it contains $\boxdot\mathbf{A}$), and Replacement of Equal Terms (if it contains $\alpha = \beta$, then it contains $\mathbf{A} \equiv \mathbf{A}'$, whenever $\mathbf{A}$ and $\mathbf{A}'$ are formulas such that $\mathbf{A}'$ is obtained by replacing one occurrence of α in $\mathbf{A}$ by an occurrence of β). Moreover, each contains Boolean algebra, classical propositional logic, S5 for $\Box$, S5 for $\boxdot$, and the following special formulas:

(R0) $\neg$ *Real 0*,

(R1) *Real 1*,

(R2) $Real\ (\alpha \frown \beta) \equiv (Real\ \alpha \wedge Real\ \beta)$, for all α, β,

(R3) $Real\ (\alpha \smile \beta) \equiv (Real\ \alpha \vee Real\ \beta)$, for all α, β,

(IU) $Int\ \alpha \supset \Box Int\ \alpha$, for all α

(P) $\mathbf{P} \supset \boxdot\mathbf{P}$, for all $\mathbf{P}$,

(UP) $\Box\boxdot\mathbf{A} \equiv \boxdot\Box\mathbf{A}$, for all $\mathbf{A}$.

In addition $\underline{L}^*$ alone contains

(I) $(Int\ \alpha \wedge Int\ \beta) \supset Int\ (\alpha \frown \beta)$.

Notice that $\mathbf{P}$ in (P) is required to be an atomic formula. In fact, even though both $\underline{L}$ and $\underline{L}^*$ are closed under substitution of terms for atomic terms, neither is closed under substitution of formulas for atomic formulas.

It is not known whether these axiomatizations are complete. (If they are, they can of course be simplified.)

4. GOLDMAN'S CHALLENGE

The puzzles and questions collected by Alvin I. Goldman in the first chapter of his book _A Theory of Human Action_ constitute a challenge for any theory of action. Ours has not yet been developed to the point where it can take on this challenge in full. Nevertheless, in this section we will try it on Goldman's curtain raiser, perhaps the least demanding of his examples:

> What is an act? One of the problems concerning the nature of acts is the problem of individuation. Suppose that John does each of the following things (all at the same time): (1) he moves his hand, (2) he frightens away a fly, (3) he moves his queen to king-knight-seven, (4) he checkmates his opponent, (5) he gives his opponent a heart attack, and (6) he wins his first chess game ever. Has John here performed _six_ acts? Or has he only performed _one_ act, of which six different descriptions have been given?
>
> ([2], p. 1)

We will first analyse this example by retelling it in our own terms (and, partly, elaborate it). Then we will offer some comments on the analysis. Finally, we will make some remarks on the strengths and weaknesses of our theory. We have collected the items under three headings: analysis, comments, and conclusion.

4.1. Analysis

John plays white, and he intends to (try to) win. The trying-to is a complication, but if we disregard that and stick to our convention of identifying intentions with (the bringing about of) events, then we may say that one of John's intentions is

> _A_ = the event of Black's being checkmated.

There is a fly crawling across the chessboard, and John intends to do something about that, too. That is, also among his intentions is

> _B_ = the event of the fly's being frightened away.

John may have many other intentions as well--standing intentions, intentions concerning simultaneous but unrelated

action, conditional intentions concerning future action--but we will think that, for the purposes of our example, _A_ and _B_ are the only ones we need to consider. Using the terminology of Section 1 we may say that _A_ and _B_ are the two intentions John has when he embarks upon his deliberation walk. Neither intention is operational. Looking at the chess board, John discovers that moving his queen _in a certain way_ will achieve _A_. Therefore he forms yet another intention:

> _C_ = the event of the White queen's moving to KKn7.

Next John finds that to achieve both _B_ and _C_ it will be sufficient to move his hand _in a certain way_. This step is such an inconspicuous, everyday kind of step that, unlike hitting upon _C_, it cannot really be said to be the outcome of conscious deliberation. However, for the sake of completeness we record that he forms a final intention:

> _D_ = the event of his right hand moving in such and such a way.

Our description of _D_ is problematic, but let us accept this for the moment and return to it under 4.2 below.

John has now arrived at an intention set _S_ = {_A_,_B_,_C_,_D_}. Suddenly there is no need to extend the deliberation walk further, for _D_ satisfies the two important conditions mentioned in Section 1: it is operational (that is, John has a routine associated with _D_), and it has the infimum property. Thus John simply runs his _D_-routine, and this is the only thing he "really does". Doing this he realizes _A_, _B_, _C_, and _D_, as well as many other events which he has not intended. Among the latter are:

> _E_ = the event of his opponent's having a heart attack,

> _F_ = the event of winning his first chess game ever.

4.2. Comments

First let us note a difference between our account and Goldman's: in Goldman's version it is not clear which things John does intentionally. It is true that the story would hardly make sense unless (1) and (3) are intentional, and John would be a poor chess player if (4) is not also intentional. The others, though, would seem open to interpretation. In our version we have cast John as doing (2) intentionally but (5) and (6) unintentionally. If one would prefer to regard the

frightening away of the fly as an unintended by-product of what John does, then just exclude B from the intention set. If one thinks John intentionally gave his opponent a heart attack, then just include E in the intention set. There is a peculiarity about F as it is logically impossible for John to realize A without also realizing F. It would make no difference to John's deliberation walk if F had been included in his intention set (given various plausible assumptions, for example, that John was only playing one chess game at a time, and this game was to be played to an end before another game could be begun). Yet it seems consistent to assume that F was not a member of the intention set.

Next let us return to the problematic description of D. In the discussion leading up to the description of B we met the underscored phrase 'in a certain way'. Thanks to the relative simplicity of chess, it was easy to give a description of B in which this phrase was given by an exact explication. A notational system has been devised which made this task trivial, but even without this invention the task would have been unproblematic since a chess move is uniquely defined by prior and posterior squares, to borrow from our own terminology. In the case of intention D, however, the situation is utterly different. Here it would be extremely difficult, if not impossible, to specify in exact terms (say, physical or physiological) just how to explicate the phrase 'in a certain way'. Therefore the equally imprecise 'in such and such a way'.

This problem brings out a difficulty which we have glossed over so far, namely, the relationship between intentions and routines. Let us go over this territory once more, more carefully this time. Agents have skills, almost all of which are the result of learning. We spend our lives, not just childhood and youth, acquiring skills--some we lose again, but new ones are constantly developed and refined. In this paper we have called some of these skills routines, without trying to characterize them further. Perhaps the main thing about them is that they are well-defined and, somehow, constructive: routines can be _run_ by the agent, and the running of a routine influences the state of the world. The set of an agent's routines is his _repertoire_. Intentions, we have said, are functions which assign to each state a set of states. Semantically speaking, routines are the same sort of thing: if the routine R is run in state w, then there is (in theory) a definite set R(w) of all the possible posterior states. Perhaps we can say that the agent considers (knows, believes, hopes, fears, etc.) that the actual posterior state will be an element of R(w), and that, furthermore, he considers, in the same sense, that R(w) is the smallest set he can define by running a routine. It may be that R and R(w) cannot be named, not expressed in the object language; that is, that

there is no term a such that $\|a\| = \underline{R}$ or that $\|a\|(\underline{w}) = \underline{R}(\underline{w})$. But if $\underline{I}$ is any intention, then $\underline{I}$ and $\underline{R}$ can be compared. We might say that $\underline{R}$ matches $\underline{I}$ in $\underline{w}$ if $\underline{I}(\underline{w}) = \underline{R}(\underline{w})$. But in general it is not excluded that $\underline{I}(\underline{w}) \cap \underline{R}(\underline{w})$, $\underline{I}(\underline{w}) - \underline{R}(\underline{w})$, and $\underline{R}(\underline{w}) - \underline{I}(\underline{w})$, are non-empty. Depending on $\underline{I}(\underline{w})$, $\underline{R}(\underline{w})$, and the agent's valuation the match between $\underline{I}$ and $\underline{R}$ in $\underline{w}$ is better or worse. Here we must leave it vague how routines are associated with intentions. But perhaps the preceding definition of an operational intention as one with which there is a routine associated should be replaced by the following definition: an intention is operational if there is a unique best routine associated with it--one which matches the intention better than any other routine, and also so well that the match is acceptable to the agent.

This perspective offers a different view of what we have called deliberation walks. The agent commits himself to one intention after the other, on the one hand trying to stick to what is essential about his original intention(s), and on the other hand trying to arrive at a sufficiently specific intention that is well matched by some routine in his repertoire. As long as there are at least two routines to choose between, he will refine his intention in order to eliminate one of the routines (or else how would he know which one to run?). At the end of a successful deliberation walk the agent is left with a unique best routine--one that approximates what he intends in what, all things considered, is for him an optimal way.

Now back to John's hand movement. Normal adults are experts at moving their hands, and it seems plausible to assume that, given his perception of the situation, John has an understanding of what $\underline{D}$ is, as well as a best routine associated with it. That we--and John--are hard pressed to give a really precise description of $\underline{D}$ is another matter. One would often identify an action of this kind by mentioning ends (that it should get the White Queen to KKn7 and leave it there, that it should frighten the fly away); one might also mention normalizing boundary conditions (that it should be a comfortable sort of movement, that it should not take too long to perform). Such conditions, though, are only necessary, not, in general, sufficient. Put another way, what is it that characterizes an element $\underline{x}$ of $\underline{D}(\underline{w}_0)$ given that $\underline{w}_0$ is the prior state of John's action)? Certainly in $\underline{x}$ the White Queen is at KKn7 (and all the other pieces as in $\underline{w}_0$), and the fly is not on the chess board. Moreover, the world must have gone from $\underline{w}_0$ to $\underline{x}$ in an "acceptable" way, but this it is hard to express within the present framework (cf. below). There may also be further conditions that have nothing to do with $\underline{B}$ and $\underline{C}$, for John may have selected $\underline{D}$ with an eye to $\underline{D}(\underline{w}_0)$ and what routines he has available.

But even if we get an exact description of $\underline{D}$--and there is no reason why an operational intention should not be describable--this would not be a description of the routine called up by John to execute $\underline{D}$. In fact, let $\underline{R}_D$ be John's routine for $\underline{D}$. Then, even though we have $\underline{D}(\underline{w}_0) \subseteq \underline{B}(\underline{w}_0) \cap \underline{C}(\underline{w}_0)$, it may well happen that $\underline{R}_D(\underline{w}_0) \nsubseteq \underline{B}(\underline{w}_0)$ or $\underline{R}_D(\underline{w}_0) \nsubseteq \underline{C}(\underline{w}_0)$.

4.3. Conclusion

Our analysis of Goldman's example registers both success and failure. The success lies mainly in the fact that we have been able to model Goldman's example in a reasonably faithful manner, without sacrificing formal rigour. Perhaps it is also fair to say that our analysis helps one to see why some authors, such as Anscombe and Davidson, have wanted to say that John "did just one thing" (the one-action view), while others, for example, Goldman himself, have maintained that John "did several things" (the many-actions view). The many-actions view is supported by some features of our analysis: John acted on a set of several distinct intentions, and he realized a number of events, some of which were intended by him. But other features support the one-action view: a unique intention was operational, a unique routine was run, a unique total change was wrought by John's action.

Causal and epistemic concepts are missing in our theory, and perhaps this is its most notable defect. Adding such concepts is a major task. A defect much easier to remedy is the imperfect concept of change that goes with our present theory. Our theory in effect identifies change with pairs of prior and posterior states. This is a promising beginning but not enough, a point made by Davidson in criticism of von Wright ([1], p. 113). Given that there is always a prior state and usually a posterior one, why should there not be some in between as well? In fact, if the actions analysed are not in some strong sense atomic, then intermediate states will exist. Thus what we need is a change function which does not pick just one posterior state but rather a <u>sequence</u> of states, the last one of which is the posterior state.

If our theory is improved by the addition of causal and perhaps epistemic concepts, and the change function is upgraded, then we will be in a better situation to meet the full force of Goldman's challenge. Among inviting tasks for future work are (i) to review concepts of agency and ability along the lines suggested in [6]; (ii) to give a systematic account of what Goldman calls level generation; and (iii) to provide for sequential action. For (iii) recent work done in dynamic and process logic may be useful; see [3], [4], and [5].

University of Auckland

REFERENCES

[1] Davidson, D.: 1980, Essays on Actions and Events, Clarendon Press, Oxford.

[2] Goldman, A.I.: 1970, A Theory of Human Action, Prentice-Hall, Englewood Cliffs, N.J.

[3] Harel, D., Kozen, D. and R. Parikh: 1980, 'Process Logic: Expressiveness, Decidability, Completeness', Research report, April.

[4] Pratt, V.R.: 1979, 'Process Logic', Proceedings of the 6th Annual A.C.M. Symposium on Principles of Programming Languages, pp. 93-100.

[5] Pratt, V.R.: 1980, 'Application of Modal Logic to Programming', Studia Logica 39, 257-274.

[6] Segerberg, K.: 1981, 'Action-Games', Acta Philosophica Fennica 32, 220-231.

[7] Segerberg, K.: forthcoming, 'The Logic of Deliberate Action', Journal of Philosophical Logic.

[8] von Wright, G.H.: 1963, Norm and Action, Routledge and Kegan Paul, London.

Kalidas Bhattacharya

SOME PROBLEMS CONCERNING MEANING

1.

The two terms 'The Morning Star' and 'The Evening Star' denote one and the same thing and yet in the very act of denoting it--one may say, for that denoting purpose itself--describe it in two different ways. This is such a simple phenomenon that none but philosophers could get interested in it. May not a particular thing, or even a particular class of things, have two or more distinguishing features, and may we not distinguish it, specify it or point to it through any one of these features, 'through' meaning in most cases stating the feature concerned in so many words? These features may, in their turn, be complex, i.e. breakable into simpler constituents; but, broken that way or not, they distinguish the thing only as they are taken, in each case, in that complex form. May be, at some incautious moments some unnecessary characters creep unnoticed into the complex. But that is never a problem: sooner or later one will have to rid oneself of them, and mostly when one resorts to analysis one does it precisely for that kind of purge.

Nor do we, in our day-to-day life, complain against the so-called indiscriminate use of terms like 'denote', 'describe', 'sense', 'meaning', or 'reference'. Every time in the specific contexts in which these terms are spoken we quite understand them, and when they sound ambiguous a little questioning here or some sifting there makes the speech intelligible. Why should one be so fastidious about precision in linguistic expression unless one sets about purely mechanical computation?

In common parlance we never also bother about what it is that means, refers, etc., whether it is the word itself or the speaker/hearer that uses[1] the word? For common people each alternative is as good as the other. They know--instinctively, if you like--that the words and their meanings, which both they have learnt from others, mostly from the seniors of their society, are but those which, equally, were used by those others exactly with those meanings. Those others too had learnt the same words and their meanings from their elders in turn, and so on, which explains how a traditional (conventional) relation, continued for a sufficiently long time, passes for an objective situation over there. Toward the

B. K. Matilal and J. L. Shaw (eds.), Analytical Philosophy in Comparative Perspective, 173–187.

beginning of any such convention one could indeed be aware that it was all a convention, a prescription--one might even say, a sort of imposition, though with no dearth of bona fides because only the respected wise men of the society were at the helm of affairs. Yet, granting all this, the stark truth is there, viz. that (i) in most such cases the origin remains untraced (even untraceable) and (ii), where traced or traceable, it is soon forgotten or even deliberately ignored, the whole thing having only a historical flavour without any epistemological import. The result, in either case, is that the relation between the symbol and the symbolized comes to be treated as a full objective feature. 'Comes to be treated as a full objective feature' implies, however, that its subjective origin, which subjectivity somehow clings to it throughout and just for which reason the relation in question is still called convention, is not at least at any stage denied. It is a sort of amalgam of subjectivity and object. What precisely such amalgam could be we shall soon see. But of two things we are assured by now:

(i) The relation between the symbol and the symbolized here, i.e. where the symbol is a word or a combination of words--may be, even a whole sentence, if not in some cases a full paragraph--and the symbolized some object or objective situation over there, is not wholly subjective, not felt as constructed or manipulable by me as I like. There is some sort of compulsion somehow imposed. The compulsion felt is not also all sham, nor simply a matter of habit. The relation between the symbol and the symbolized may have been, as it is often found to be, originally prescribed by wise men or men of power whom we feel obliged to obey. But, first, as we have just seen, such prescriptive cases are only a few compared to the vast number of other cases where the origin is so much forgotten that there is no trace left for reconstructing the whole show. Secondly, even in cases the relation is definitely known to have been prescribed by men we feel obliged to obey, this our feeling so obliged contributes a lot toward the objectivity of the relation. As long as we feel so obliged and do not at the same time feel that we are obliging those esteemed men, we, to that extent, recognize a sort of objectivity in their prescriptions, however irrational these prescriptions may otherwise be. This is the principle behind what is called śabdapramāṇa in our Indian philosophy. Thirdly--and this and what immediately follows are the most important considerations here--feeling obliged this way is a sort of sanctification, however temporary or irrational,[2] and sanctification of a person is not exhausted in my treating him in a particular way, but finding in him virtues like sooth-seeing, sooth-saying, etc. True, this my knowledge of him as a sooth-seer, sooth-sayer, etc. may prove wrong as any other preliminary knowledge may prove to be. But, then, this

contingency being always there, our motto should be: we abide by whatever knowledge we have as long as it is not contradicted or replaced but reject it as soon as we meet with such contradiction or replacement. 'Knowledge' here does not indeed stand for any idea (of an object) that we happen to have: it, as distinct from belief, must have something additional to speak in its favour. The favouring consideration in our present case is the sanctity of the person concerned, as distinct from any other lay man. This sanctity <u>may</u> at any time be blown off by the slightest misdemeanour of the person concerned or for any other reason. But the simplest logic is that we should not only not hesitate about him till any such consideration is forthcoming but also, as in every other case, go on defending him as long as possible, on condition, however, that when all these attempts fail we should not hesitate for a moment to discard him. This is the logic of normal human life, and this logic certifies that the relation between words and things meant is not all subjective.

(ii) Equally, we argue that the relation between words and the objects meant is not all objective either. True, quite in the largest number of cases such relations are not traceable to actual prescription by respected men of the community. In such cases, therefore, it might appear that there is no touch of subjectivity anywhere in the history of that word vis-à-vis its relation to the appropriate object. But this would be much too revolutionary just now. What merit it has will, of course, be elaborated later. For the time being we only note that the relation in question cannot so easily pass as wholly objective. It is never objective in the full sense in which one understands the relation between a cause and its effect. A symbol being presented one does indeed, and often unerringly too, expect the symbolized, but that happens <u>only because</u> one has been trained that way, whereas in the cause-effect case this training itself has been possible because of the intrinisic objectivity of the relation, for which reason this latter training is all passive and automatic from the beginning. Indeed, it is no training at all, it is nothing more than a habit formed all mechanically.

Full objectivity of A is no other than the simple fact that A is a gross item of nature--physical or mental--directly presentable in all its grossness whether to external perception or to introspection. The word-object relation is never objective in that sense, even where the object side is gross. The word, indeed, as primarily a sound or a body of sounds, is one such gross object, and even the object side is often gross in that sense, but the meaning relation that connects the two is never so gross, never amenable to sense-perception. It is a sheer convention necessary for understanding others speaking, and also for communication and facilitation of thought in different dimensions. As such a convention it is rather on

the same footing with intellectual theories in their relation to the gross natural presentations they are constructed to explain, simplify, etc., except that while these theories are constructed on intelligible grounds the meaning relation is not constructed that way. Theories are constructed on the basis of (perceived = passively received) facts and logic (that includes mathematics). This explains why in their case there is sort of compulsion, intellectual nonetheless, from the beginning. What constitutes the primary objectivity of these theories is this their compulsive character which, in so far as it is determined by (facts and) logic, is given the generous name 'necessity'. On the other hand, the objectivity of the meaning relation, its compulsiveness, is constituted wholly of its being a convention that, unchallenged till now, has stood the test of time and is accepted on the testimony of generations of people. As for other ways of testing the objectivity = acceptability of theories, these are equally, though more crudely and commonsensely, needed in the case of meaning relations. Like theories, therefore, meaning relations too are as much objective as subjective. In neither case does subjectivity and objectivity stand in each other's way.

2.

We began with saying that healthy commonsense is never concerned with the subtle distinctions that philosophers draw between sense, reference, meaning, denote, etc., nor with the question whether it is we or the words themselves that mean objects outside, we through the words and the words of their own right. Some elaboration of these two points is needed in spite of whatever we say against philosophers.

Everything in the world has so many distinguishing features. When we refer to a thing we turn the hearer's (including our own) attention to it through any of these features, meaning we understand it, on that occasion, *qua* possessing that feature. If the thing is *meant*, i.e. constitutes the meaning of the word concerned, the feature in question is the objective *determinant of that meaning* for that occasion. Not that it is not itself unmeant altogether. The relation between it and what we call meaning (meant) is like that between means and end. If we will an end we will the means too, but whereas the means is willed as a sort of determinant the end willed is pictured as complete and determined. So is the case with meaning and its determinant--in the Nyāya language, *śakya* and *śakyatāvacchedaka*; in the good old language of J.S. Mill, denotation and connotation. Take, for example, the term 'man'. Its meaning (*śakya*) i.e. denotation, consists indeed of individual men--Ram, Rahim, David and so many others. But note that none of them constitute the

meaning or denotation on their own account. On their own account they are only items over there. They constitute the meaning of the term 'man' just in so far as they are considered in the light of, i.e. qua possessing, one or other of the distinguishing features of man, viz., rationality, ability to speak, laugh, use instruments as instruments, etc. etc. These latter constitute the connotation, śakyatāvacchedaka, of the term 'man'. So also in the case of the notorious Morning Star and Evening Star. The particular star called by both these names is the denotation of each of the terms, and Morning-Star-hood and Evening-Star-hood are in either case the śakyatāvacchedaka, the determinant. If we call them connotation, as we must, the only point we like to add to Mill's thesis is that the connotation as constituting the distinguishing mark may be different in different contexts, particularly when the object to distinguish, i.e. the denotation, is numerically only one, as in the case of the Morning Star and the Evening Star.[3]

The distinction between śakya and śakyatāvacchedaka is closely and most profitably noticeable when the term in question includes a demonstrative adjective like 'this' or 'that'. Thisness or thatness is certainly a determinant in every such case, the particular thing meant being precisely the one to which the speaker just draws the attention of the hearer through the word 'this' or 'that'. The special point to note about 'this' or 'that' is that, though thisness or thatness is a determinant, a connotation, it is no property or feature of the thing concerned. The speaker in saying 'this' just points to it as is done by a finger-post and there ends everything. Except where a class of things is meant or where even a definite particular thing is meant through some superlative or through some additional property as in the case of 'this red flower', the speaker need not intend a property.[4] Mere 'this' has nothing to do with a property of the thing meant, though after the attention of the hearer is drawn the speaker may go on adding properties. What much we insist on here is that determinants may not be, and therefore are not as such, properties of the things meant.[5]

But then we are landed in a serious difficulty. Demonstrative words like 'this', 'that', etc. are intelligible primarily as spoken and only secondarily as heard. 'This' means the thing as so spoken, i.e. pointed out by the speaker, maybe along with some gesture, and so with all such demonstratives; and when the hearer comes to know the thing he understands it necessarily as what was spoken (of) to him as 'this', not necessarily what he (the hearer) speaks (of) as 'this'. The hearer is in a perfect realistic attitude: he comes to know a thing over there through its perfect individuality indicated by the word 'this' spoken out to him. So far he does not himself use (speak) the word. The heard word means the real

thing over there in front of the speaker. Even if the thing is understood as <u>necessarily</u> in front of the speaker, i.e. in some necessary relation to him, even then this is no idealism. For the speaker, so far, is only a 'he' to the hearer, standing practically at the same level with other things of the world, the only difference being that 'he' is a living, conscious item. To the speaker, however, 'this' is related in an entirely different way. No doubt, it is as much in necessary relation to the speaker as to the hearer; but the speaker here is 'I'[6] and 'I' is no item of the world coordinate, bracketed, with the things of the world. 'I' is not just a living conscious item. It is this novelty of 'I' which inevitably distinguishes whatever is necessarily related to it not only from all that is over there independent of it but also from whatever stands necessarily related to a 'he'. Necessary relation of anything to a 'he' is every bit on the same footing with any necessary relation between X and Y where X and Y are neither persons nor even animals. This fundamental distinction between 'I' (speaker) on the one hand and 'he' (hearer) and 'it' on the other is so patent that whereas the dependence on 'I' is everywhere taken as the bedrock of idealism, dependence on 'he'[7] is not considered that way. If between any two things there is necessary relation, such that one particularly of these cannot subsist without the other, this has nothing to do with idealism; and so too is the case with the necessary relation between 'he' and 'this'. While for the speaker 'this' is not only understood as but is really in necessary relation to his 'I', for the hearer it is in necessary relation with someone he calls 'he', i.e. with the speaker who is 'he' to him. The hearer says that the 'this' in question is what the speaker called 'this'.

Normally, in philosophy the two attitudes--one of speaking and the other of hearing--are not distinguished. In western philosophy the meaning of a verbal expression is most often, if not always, understood from the point of view of the speaker, as what the speaker means when he uses that expression. In classical <u>systematic</u> philosophy in India, on the other hand, it is almost always--the only exception being Buddhism--understood from the point of view of the hearer. This is why in epistemology Indian systematic philosophers have, with the exception of the Buddhists, been realists uniformly.

True, western philosophers, understanding language and meaning mostly from the side of the speaker, have not yet always opted for idealism. I am not here speaking anything for or against that form of idealism for which sense-data are nothing but sensations. I am referring here to the so-called higher type of idealism, often called transcendental idealism (as distinct from empirical, psychological, idealism) for which thought constitutes the structure of Nature down to the detailedmost, viz. the generic features (i.e. universals) like

cowhood, horsehood, etc. That the structure (in a hierarch of detailedness) is constituted by 'I'--in Kant's language, by 'I think' equal to, one might say, 'I speak'--first, suggests itself incontrovertibly when a speaker uses the word 'this'. Not that the thing the speaker indicates by 'this' is reducible wholly to this his speaking 'this', i.e. pointing. Some thing indeed is pointed out, but what immediately interests the speaker is its 'thisness' which is the determinant here. In other words, the thing is of interest just so far as it is 'this'. Once that interest is satiated, both the speaker and the hearer may get interested in further details. But, to start with, these all stand engulfed in that spoken 'this': they are all anticipated in a most general manner--one may even say with Husserl, bracketed though touching all the while and all along the periphery--another way of saying which is that _I_ refers to all this _sui generis_. My knowledge at least at the stage of 'this'--in other words, my speaking 'this'--is of its very constitution so referring, so forward-looking, that like a teased angry cat ready to pounce on any living thing that passes its way, it, constitutionally ready to refer, refers actually to whatever comes on the way. 'This' is but a sort of crying out such intrinsic constitutional reference by 'I' the speaker. Genuine 'this', 'this' as spoken, not heard, is the basic _a priori_ anticipation, in this way, of anything in Nature, and there is not much of distinction here between any thing and all things. The _word_ 'this' is the clearest case where a word coalesces (idealistically, of course) with a presentable thing of Nature _qua_ presentable. This is the basis of a mystical doctrine prevalent in India, viz. that a name and that which it is a name of are identical. This doctrine, if true at all, would obviously be true of the relation between the act of naming and the _most basic_ characters of the thing named.

Next in concreteness, though still very largely _sui generis_ in reference, is the idea of the common feature of a class of things, called universal. This will be clear from how we know the universal. Unless a universal is foisted on us _ab extra_ through a sort of theory-construction, we never speak of it until we have perceived at least one instance--one cow at least when the universal is cow-ness. But the beauty of the whole thing, scarcely noticed by philosophers, is that one instance experienced is enough. The larger the number of instances the better the universal character may come to be defined through abstraction. But even of this we are not always assured. What much is beyond dispute is that if a child has seen one elephant, say, in a picture, and is told that it is called elephant, the very next time he perceives another--actually or in a picture--he recognizes it and speaks out 'elephant'. He had no time to compare the two instances and then abstract out the common feature(s). Immediately as

he saw the second instance he recognized it as an elephant. This he could not have done if he had not directly perceived the universal Elephant in the first presentation. Yet to say this is not easy enough, in spite of all the intellectual eloquence of the Naiyāyikas. For, if it is held that in the first instance the child has perceived the universal Elephant it will have to be equally held that in the same instance he had perceived so many other universals--indeed, a whole host of them arranged in a hierarchy. Not that this is theoretically impossible. But there is an alternative simpler solution in easy consonance with other activities of the mind. It is that on experiencing any thing our mind feels constitutionally disposed to anticipate similar things--in plain language, remains ever on the look out for similar things--all sui generis, and immediately as any such similar thing appears on the scene recognizes it. What the name 'elephant' means is not this or that particular elephant, nor even all elephants considered collectively, but just any elephant, any animal expected to fit in with the generic image of elephant; and hence whenever such an animal passes one's way one recognizes it. Considered thiswise, it stands on the same footing with what is meant by 'this', only more empirically loaded. All this, however, for the speaker only who calls the presentation 'elephant', not for the hearer who only learns from outside that the name 'elephant' is associated with such-and-such animal and expects nothing sui generis. It is from the speaker's point of view that, regarding the universal, we have, as much in Indian as in western philosophy, theories like realism, conceptualism and nominalism. The name coalesces with--totally embraces--that which it is a name of, as we found in the case of spoken 'this'. This exactly is, in the case of the universal, called nominalism. As there is such coalescence, the whole situation can also be understood as much to be my speaking, i.e. anticipating, similar presentations sui generis as also the objective generic feature, the universal, over there. The former of these two is conceptualism and the latter realism.

Once the word 'this' and a class-name are understood this way, it is not difficult to understand the terms 'space', 'time' and those standing for the Kantian and post-Kantian categories and the phenomenologists' 'essences' in a more or less similar fashion. Indeed, if 'this' stands for an anticipated particular as just a presentation and other words like 'cow', 'elephant', 'red', 'green', etc. for anticipated--and all such anticipation is sui generis--empirical contents, it is not difficult to see that space, time and the Kantian and post-Kantian categories are what are so anticipated by words speaking syntactical relations among such substantive words. All this, however, from the speaker's point of view. There is nothing like this for the hearer's 'I'. The hearer can at

most know--'perceive' if you like, 'theorize about' if you like--what the situation is or was for the speaker. A word meaning a thing is thus really the speaker's meaning, which is another way of saying that a subject, some genuine 'I', means the thing through that word; and that word may be as much substantive, standing for some thing or attribute, as syntactical, and equally in very many cases as both, i.e. in the form of a whole phrase, short or long, or even as a full sentence (sometimes as full even as a paragraph).

Before we pass on to the next section, a reminder once again. Although we have till now used the terms 'mean', 'denote', 'indicate', 'refer' etc. indiscriminately we feel sure we have not talked nonsense. Wherever there was the least chance of being really misunderstood we took precaution and clarified the use of the terms concerned, as was the case, for example, with 'connotation', 'determinant', etc.

The type of philosophy we presented, though in outline, in this section was not only the transcendental philosophy, sometimes dismissed as speculative metaphysics, of the West; it was also in rudiment the logos philosophy--the _śābdika darśana_ of some Grammarians, the Śaivas and even of some Buddhists--that developed in India but was ignored, almost literally, in the Indian philosophical _systems_. Ignored, because while these philosophers understood words, including phrases and sentences, mostly if not always, from the speaker's point of view the systematic Indian philosophers quietly passed on to that of the hearer.

One thing, obvious enough and therefore not separately attended to till now, is that speaking and thinking are one and the same activity, except that when a hearer, as hearer, thinks, i.e. _understands_, this is only his rehearsing a possible speaker's original thinking, always retrospective and never originally forward-looking. It is on this interpretation of thought, viz. that thought and speaking are the same activity, that we brought in Kant, post-Kantians and the phenomenologists to our support.

3.

In the preceding section we held that if substantive words mean things and qualities each in its isolated individuality, a whole phrase, sentence or paragraph means a sort of complex affair composed as much of these isolated individuals as also of diverse relations between them, including the overt activities meant by certain verbs.[8]

An important question which has, in this connexion, troubled all linguistic philosophers, both in India and the West, is whether the meaning of a sentence[9] is some sort of later aggregation of the meanings of the individual words that compose it or whether the whole meaning has somehow been

operating throughout, getting only clearer and clearer, more and more defined, through the separate meanings of the individual words and that too, in the latter case, in diverse measures according as the contributions of the meanings of the constitutive words differ in importance. The first theory is known as brick-and-mortar and the second as holistic. In the West the debate started in right earnest only in recent days with the analytical thinkers starting the ball rolling and others of more or less transcendental temper (loosely called speculative, metaphysical etc.) sending it back to their court. The transcendentalists argue that the whole, as everywhere, is functioning all through starting as incohate, i.e. pointing, from the beginning to an as-yet-incohate whole, but getting more and more explicit, defined and full-bodied through the single words getting more and more related to one another.

In India this debate had continued for ages, particularly with the Naiyāyikas on the one side and the Mīmāṁsakas on the other, the Mīmāṁsakas differing among themselves in matters of detail. The Naiyāyikas as confirmed realists, pluralists and analysts have very naturally opted for the brick-and-mortar theory. They hold that the individual meanings of the constitutive words have first to be learnt and then combined with one another according to their position in the sentence heard, but moderated by the natural affinity of some classes of words with some other classes and also in some cases (particularly where literal meaning is for any reason unacceptable) by what the hearer wants to know. One who is by open admission an empiricist, a pluralist and an analyst, need not deny relations (even in some cases necessary relations) among constituent units and the formation of wholes accordingly. He need not even deny that such wholes, once formed, maintain independent existence of their own till they get dissipated once again into those units. Only, according to him the constituent units come first then the wholes formed. He is not in the least unnerved by the neo-Hegelian type of argument which, in India, he had to meet much earlier, viz. that relations to relate data require further relations to relate themselves to those data, which according to those neo-Hegelians (in our country, the Advaita-Vedāntins like Śaṁkara) is unacceptable because of infinite regress involved. Much like a modern pluralist he holds that relations proper are reals of a different category altogether. He only adds that not all reals which ordinarily pass as relations are so categorically different, they requiring further relations to get related to data. The final relation that relates two reals does this, he holds, by itself walking into either or both of the relata and that way getting identified with them. We need not work this out any further. The Naiyāyika has developed his account of language and meaning as consistently as possible with his

empiricism and pluralism, interpreting the meaning of every sentence as developed out of the meanings of the constituent terms got related to one another according to his theory of relation.

Nyāya is consistent in another procedure too. Language it understands throughout as heard, not as spoken, the speaker being understood as only an erstwhile hearer learning language and its meaning from other speakers and now only using them. [True, these other speakers had in the same way learnt language through hearing and, as speakers, had only used the same. This, the Naiyāyika admits, leads no doubt to indefinite regress. But he is not unnerved by that: where such a regress is a fact one should not speak ill of it.] We have already seen that from the point of view of hearing there is no sui generis anticipation of a whole, no a priori unity lending structural relations to the constituent words and their isolated meanings. This is done, we have seen, by one who speaks, not by one who learns through hearing. Hence the Nyāya has proceeded consistently all along (though that does not mean that its procedure has been correct).

But the whole thing appears anomalous when we turn to the Mīmāṁsā and the different schools of systematic Vedānta. They all, like Nyāya, consider language from the hearer's point of view and yet admit anticipated wholes which are intelligible only from the speaker's point of view. Then, again, Mīmāṁsā is as empirical and pluralistic in attitude as Nyāya. So, for it to have insisted on anticipated wholes looks doubly anomalous. All systematic Vedāntists are transcendentalists. Their transcendentalism justifies indeed their holistic attitude, but what about their understanding language always from the hearer's point of view? Perhaps in their case, as much as with Mīmāṁsā, philosophy as transcendental spiritual exercise has somehow continued in their systematic intellectual account and made its presence felt.

Anyway, even from the hearer's point of view Mīmāṁsā has insisted that the minimum living unit of language is a sentence, however short, ending with a verb. Rather, according to it, there can be no learning a word unless it is verb-centric. The Neo-Hegelians in their days in the West, particularly Bosanquet, developed such holistic logic so verb-centrically.[10] Some other Mīmāṁsakas, however, advocated holism without any special insistence on verbs. Their point--and that is also the point of many Vedāntins--is simply that we never understand the meaning of a single word unconnected with some other words, and also that from the side of the reals meant we always understand these reals, in every particular context, not only as just connected with one another but as forming a genuine whole. The genuine whole, whether of meanings or of words, they hold, was there, prior to its complete manifestation. It was originally only

implicit, getting gradually explicit and full-bodied according as the so-called elements showed themselves there. This in its most self-consistent form is the old satkārya-vāda of Sāṁkhya and of the Vaiṣṇava and Śaiva monists, and the vivartavāda of the Advaita-Vedāntin was just an advance upon it in a certain special interest. But why the Mīmāṁsakas and other pluralists opted for such a view is not easy to understand. The only way we could understand this for pluralists (and realists) is to approach the whole situation from the pòint of view of speaking, hearing always being in retrospect. Only then we could say that even for pluralists (and realists) a genuine whole is anticipated sui generis, though yet unformed in detail, more and more details appearing only according as the anticipation, still sui generis, turns toward more and more detailed contents. Another name of such anticipation is speaking = thinking (which is not in retrospect, i.e. no understanding). But none of the thinkers under consideration, except the Śaivas, Buddhists and some Vaiṣṇavas, have that way insisted on speaking, rather than on hearing. Systematic Indian philosophy (except Buddhism) has, as we have already said, insisted almost cent per cent on hearing, i.e. on knowledge as passive happening rather than as active manipulation. This is such an obsession with them that they have understood even inference as principally what just happens, like Hume's associational passage. Whatever activity, called tarka, may be involved in inference and may sometimes rebel against passive associational acceptance is promptly snubbed as off the normal life of knowledge; and even for the spiritual life of sādhana such passivity is held in equal esteem by them, all sādhana starting, according to them, with unquestioning acceptance and ending with the Ideal just presenting itself--passivity as much at the start as at the end. Though all sorts of questioning and manipulation, intellectual or yogic, are permitted midway between the beginning and the end they are deliberately subordinated to and submerged in over-all passivity--an attitude quite in tune with hearing esteemed over speaking. [Unsystematic Indian philosophy, like Śaivism, other forms of Tāntrism and some forms of Vaiṣṅvism, however, is free from this limitation.]

This anomaly we find equally in western philosophy. Western people have for long ages been treating language as primarily speaking, not hearing, and thinking as primarily 'I thinking', i.e. anticipating sui generis thing structurally, not as understanding, i.e. re-constructing them, consciously rehearsing what an original thinker would have done. Consistently with this speaking attitude as dominating over hearing, they all ought to have advocated the holistic attitude and preferred transcendentalism to empiricism. But for the last four hundred years the number of empiricists and pluralists among them, which only the hearing attitude could produce, has

been steadily increasing,[11] pushing away all forms of transcendentalism and holism. This vicarious growth there is traceable to a contradiction systematically nurtured in western life during the last four centuries. Though constitutionally in the speaking, asserting, dominating attitude, i.e. much more on the side of freedom than for passive acceptance of the given, the western people have, through these centuries, employed all this freedom, all this assertion and domination over Nature, ultimately to facilitate their natural animal life, to ransack Nature in order to collect all comforts ultimately for that natural animal life. Freedom which is the prerogative of man and meant for turning Nature into a kingdom of heaven where every man--and wherever possible, animals, too--should be taught to have and exercise this freedom at its highest has rather been steadily sacrificed at the altar of Nature, at the altar, in other words, of the animal that is Nature in man. The subtleties of thought have all been developed only to understand Nature, to think about it, once it is there, in all its finest details, and then the details are made to combine mechanically into various sorts of unities. This is but considering Nature in retrospect, reconstructing it and, in the freedom of reconstruction, reconstructing it in all possible ways and then choosing the one that is best working. This is not speaking out Nature, no original anticipation of its structure, as the transcendentalists have in all ages done; it is sort of hearing it once it has been spoken out, and--what is more--a form of aggressive hearing, not only not caring for the fact that it was originally spoken out but, worse, denying that fact altogether, denying even its possibility, castigating that way all transcendental philosophy as sheer imagination (speculation) but never for a moment perceiving the truth that at least this imagination is characteristically human and therefore inescapable in spite of all the condemnations piled upon it. Understanding, too, like hearing, is indeed equally human, i.e. a form of freedom. But these moderners gloat over its being tied to Nature, freedom binding itself to what just happens to be there. Without their knowledge they have been led to such a pass that not satisfied with such givens they have struggled desperately to understand even logic, i.e. the constitutive functionality of thinking = speaking, in terms of the given, sometimes, as with J.S. Mill, as just some highest generalizations of the given (and even those generalizations as no characteristic human activities but simply as generalities somehow streaming out of the givens[12]), sometimes as conventions and to that extent objects over there, sometimes as means (postulates) but with this distinction that whereas in most cases the means are also ends, though subordinately, the means here in question ever remain means, except when in a meta-attitude they come to be considered artificially by

themselves. These people have almost religiously remained blind to the simple truth, accepted by common people even now-a-days, that the principles of logic are but the original, unsoiled, ways of thinking or speaking (in the case of speaking they are called semantic or syntactical principles, rather than logic, though at bottom the two ways of treating are ultimately the same). Error, for this simple theory, is due to the misdemeanour of the given, its disclaimer that originally it was spoken, and when the disclaimer is loud enough speaking itself gets confused giving rise to what is ordinarily called contradiction or formal error.

While medieval India, forgetting its brilliant tradition of logos philosophy,, i.e. the philosophy of speaking, turned over professedly to the hearing attitude, though through the systematic philosophies it constructed the old speaking attitude peered in the form of holistic attitude, the history of modern West has moved in just the opposite direction. Loudly eulogizing speaking, assertion, freedom or thinking, and asserting it all through, it has quietly, and even as unnoticed by them, passed on to the attitude of hearing, understanding, acceptance and submission to Nature, interpreting thought itself in terms of the given, as convention, instrument and even sometimes as highest generalities exuded by the given Nature itself.

NOTES

1. The use, as already said, is mostly in the form of stating, i.e. some linguistic expression.

2. Indeed, sanctification as such is always non-rational, and even so the sanctifications of reason itself.

3. We have called the denotation 'meaning'. Heavens would not fall had we used the terms 'sense(d)', 'refer(red) to' and the like. So long as our ideas remain clear use of different terms to indicate the same thing does not matter.

4. After what we have said _proper_ _names_ need not be considered separately.

5. Be it noted further that when a property or a class of similar properties is meant the determinant is some further specifications of that property, and in that sense, a property of the property concerned.

6. Not that the hearer cannot say "I hear 'this'", but this is a self-conscious affair far less frequent than hearing that is not so conscious and only acted up to. Further, even when self-conscious, hearing has to toe the line of speaking, speaking that way being more fundamental.

7. Considerations of 'you' as distinct from 'I' and 'he' is avoided here as it offers nothing special for or against idealism.

8. All grammatical verbs do not mean overt activity.

9. For the time being we ignore consideration of phrases and paragraphs.

10. Nyāya too has not wholly denied this. Every child, he holds, learns word-meanings normally through _sentences_ heard, but he promptly adds that when through a process of gradual selection through mutual ouster (_āvāpodvāpa_) resulting from hearing diverse sentences like 'Bring the cow', 'Drive away the cow', 'Bring the umbrella', etc. the child learns the exact meanings of the constituent words holism is gone.

11. There has occurred some serious break-through only in recent days.

12. This view is not, of course, very much in vogue nowadays.

S. Bhattacharya

ABSTRACTION, ANALYSIS AND UNIVERSALS: THE NAVYA-NYĀYA THEORY

1. INTRODUCTION

The concept of universal has been formulated in various ways to solve various types of problems, to serve various, often conflicting purposes. So a formulation specially suited to solve one type of problem, fails, for that very reason, to solve other, conflicting, types of problems. From the very nature of the case it would appear impossible to have one concept to do all of the work which is naively expected of it. To explain our points we list, without attempting to be exhaustive, nine different pairs of terms:

(1) universal --- particular
(2) abstract ---- concrete
(3) essence ----- accident
(4) form --- matter
(5) reason --- sense
(6) reality -- appearance
(7) a priori ---- a posteriori
(8) adjective --- substantive
(9) predicate --- subject

It is usual to connect, and where possible to identify, some of the first terms of the pairs with universals and the corresponding second terms with particulars. It is obvious that the first terms of all the pairs cannot be identified with universals. For example, if universals are regarded as reals and particulars as appearances (pair (6)), then it will not be possible to identify them as predicates and subjects; for, even if reality be regarded as a predicate, it will not do to have reality as a predicate and an appearance as a subject of a judgment (proposition, sentence). Moreover, the pairs belong to different spheres--logical, epistemological, grammatical, ontological, etc.--and involve different types of problems which require different methods of solution.

2. MEANING OF GENERAL TERMS AND UNIVERSALS

It is usual to introduce the concept of universal via a theory of meaning of general terms, e.g., common nouns and

B. K. Matilal and J. L. Shaw (eds.), Analytical Philosophy in Comparative Perspective, 189–202.

adjectives. It is pointed out, a common noun like 'cat' or an adjective like 'wise' means any and every one of a class of _selected_ individuals, e.g., particular cats and wise individuals. This process of selecting individuals and collecting them into a class, it is argued, remains inexplicable except on the assumption of a property (or set of properties) shared in common by all and only those individuals to which the general term applies. This argument justifies the postulation of a self-identical common property as the _ground_ for the application of a general term to all and only those individuals to which it applies. A general term applies, refers to, or denotes, all individuals which share some common properties which the term implies or connotes. (The class or collection of individuals is _not_ denoted by the general term, but only its members).

This theory of general terms leads to a particular theory of universals via a definition of 'universal':

D1. x is a universal $=_{df}$ x is present in many at the same time.

Ramsey, however, has an objection against this definition of 'universal'. He says:

> Next we have various distinctions between objects based on their relations to space and time; for instance, some objects can only be in one place at one time, others like the colour red, can be in many. Here again, in spite of the importance of the subject, I do not think we can have reached the essence of the matter. For when, for instance Dr. Whitehead says that the table is an adjective, and Mr. Johnson that it is a substantive, they are not arguing about how many places the table can be in at once, but about its logical nature.[1]

But this objection of Ramsey depends wholly upon his identification of _adjective_ with _universal_, of _substantive_ with _particular_. So instead of being an objection against D1, this may be interpreted as an objection against this identification.

Now the properties common to all the referents of a general term are present in all of them simultaneously. So the common properties which general terms imply in order to apply to whatever they apply, are universals in the sense of D1, and the referents of the general term in which the universals are present are the _instances_ of the universals. We shall call this theory of universals 'The Theory A'.

The postulation of universals as reasons for applications of general terms also explains why expressions which are grammatically general terms but which do not, because they cannot, apply to anything, cannot imply any universal. Quine's question (raised by him in a different context of empty *singular* terms) 'Having cluttered the universe with an implausible lot of unactualised possibles, are we to go on and add a realm of unactualisable impossibles?' may be raised about empty general terms too.[2] The Theory A answers this question in the negative, for universals are postulated as *grounds* or *reasons* for the application of general terms, i.e. as *necessary and sufficient conditions of their application*. If a general term fails to apply to anything at all, then there cannot be any universal implied by it; the sufficiency guarantees this. If there is a universal implied by a general term, then it applies to its instances. The apparently general term 'G' does not apply to anything; hence there is no universal implied by it. So expressions like 'spherical pyramids of Copilco' do not imply any universals.

This theory, specially the argument, has been criticised from different points of view. The most fundamental objection is by Quine. He points out that the so-called general terms are not really terms at all, but only 'constitutent syllables' "comparable to 'rat' in 'Socrates'."[3] The Theory A analyses the sentence

S1. Socrates is a man.

into three parts--two terms, 'Socrates' and 'man' and the copula 'is'. Every sentence is analysed in this fashion, so that we have always at least one general term as part of the sentence. Quine, on the other hand, analyses S1 into two parts--(1) 'Socrates' and (2) 'is a man' called predicate in a new sense. He argues, "The positions occupied by general terms have indeed no status at all in a logical grammar, for we have found that for logical purposes the predicate commends itself as the unit of analysis."[4] As there are no general terms, there is no problem of explaining how or why they can denote what they denote. The problem has vanished, and with it, the need for a solution. We shall call this theory 'The Theory Q'.

The Theory Q, however, has its own difficulties. For even if the Theory A stems from a faulty analysis of sentences, still in the Theory Q the problem of meaning of predicates takes the place of the problem of meaning of general terms. Quine himself admits the legitimacy of this new problem. Thus talking of the predicate "(8) ① used to work for the man who murdered the second husband of ①'s youngest sister", Quine says "If a meaning for these strange expressions called predicates be demanded, e.g. for (8), an answer is

'former employee of own younger sister's second husband's murderer'." But then we have got a general term as the meaning of 'predicate' in Quine's sense. So it would seem that Quine has not succeeded in eliminating general terms.

Moreover, Quine is not consistent about the radical difference between general terms and 'predicates' in his sense. Introducing the notion of predicate in his sense, he says,

> A retreat to the view of capital letters as representing terms, absolute and relative, thus seems indicated. When a capital letter occurs monadically ... we may interpret it as representing an absolute term; while it occurs dyadically ... we may interpret it as representing a dyadic relative term; and so on. Thus 'F' and 'G' ... may be explained for purposes of the example of the philosopher as representing respectively the absolute term 'philosopher' and the dyadic relation 'contradicts'.[5]

> ...circled numerals may be viewed simply as a supplementary device, more convenient and systematic than those existing in ordinary language for abstracting complex terms out of complex sentences. Thus the shift which we have made from terms to predicates can be viewed as a case of merely improving and renaming the idea of term.[6]

Thus it is not clear, whether according to Quine, a monadic predicate, as the unit of analysis, is radically different from a general (absolute) term, or is "a case of <u>merely</u> improving and renaming of the idea of term."

An objection from a different point of view has been brought against the Theory A. It has been pointed out that there may not be anything positive common to all the species under a genus or to all instances of a universal. The different colours, for example, do not have anything positive in common; rather they differ among themselves in a characteristic manner. A colour, say red, differs from all other colours in a way which is totally different from the way it differs, say, in shape. What all colours have in common is not any positive property, but a common way in which every colour differs only from every other colour.

We should note here a feature of this type of argument against the Theory A. The argument given for the Theory A is a transcendental one--it will be inexplicable otherwise how a general term can be used to refer to all the particulars to which it does refer. But the above objection is that, as a matter of fact, many general terms, like 'colour' do not imply

any positive property common to all the particulars. This is an empirical argument against the transcendental argument of the Theory A.

This type of empirical consideration to decide a philosophical issue has become very popular in recent philosophy. The exhortation 'don't think, but look' in order to find out what, if anything, is common to all instances of a universal which a general term implies in order to apply to them, highlights this type of approach to philosophical problems. But whether a philosophical problem can be solved in this way, or a philosophical theory can be refuted by empirical considerations, is an open question. Alternatively, it may be asked if the question: 'How can general terms apply to all and only those particulars to which it does apply?' is an empirical question.

There is, however, a philosophical argument behind this empirical consideration. This argument refutes the Theory A by pointing out that it is not necessary for all referents of a general term to share one identical property; it is sufficent if the different referents have 'family resemblances'. It is not necessary to have one identical thread running through all the referents, it is enough if there are overlapping fibres of varying lengths. We shall call this theory 'the Theory W'.

<u>This</u> point of difference (for other important points of difference see below), between the Theory A and the Theory W, can be logically explained by a difference in the scopes of quantifiers, thus:

The Theory A asserts: (LA) There is something common to all referents of a general term.

The Theory W asserts: (LW) All referents of a general term have something in common.

What is common is independent of the choice of referents in (LA), where as it is dependent on the choice of referents in (LW). This way of formulating the difference between the Theory A and the Theory W shows that the former is a logically stronger theory than the latter; for (LA) implies (LW), but not conversely. That is, <u>if</u> there is something common to all referents of a general term, this will explain, according to <u>both</u> the theories A and W, why the general term applies to whatever it applies to. (For further discussion of this point see below).

3. ABSTRACTION, PARTICULARITY AND UNIVERSALS

There is a definition of 'particular' corresponding to D1 of 'universal':

> D2. x is a particular $=_{df}$ x is present in a particular region of space and/or in a particular period of time.

Particulars so defined are often identified with instances of universals, and are so denoted by singular terms--proper names and definite descriptions.

If instances of a universal can only be particulars as defined in D2, then only the instances, but not the universals, are the existents. As a matter of fact, the particularity of particulars defined by their location in space and time is the same as their existence. For, to exist (it is often asserted) is to be in space and time. Thus particularity and existence of particulars become the same. A universal as defined in D1 is not a particular as defined in D2, for a universal is present in all instances, past, present, and future, at the same time. This may be taken to mean that universals transcend the limitations of space and time, or transcend space and time altogether. Thus it will not make sense to say that humanity, for example, is near or far, or has spatial or temporal dimension.

Now if instances of a universal are identified with particular in the sense of D2, then there cannot be any non-existent instance of a universal. But it is usual to make a distinction between 'square circle' and 'golden mountain' by holding that while the former <u>cannot</u> apply to anything at all, the latter like any other general term, has denotation, although it has <u>only</u> possible instances and no actual instance. But how are we to understand the phrase 'possible instances'? In the language of possible world semantics, possible instances will be instances in a possible world. Is this possible world in space and time? How is this then different from the actual world? Or, are there possible spaces and times different from actual space and time? But possible spaces, like n-dimensional space where n is greater than three or four, will be merely concepts and will not, therefore, be able to function as principles of individuation or particularisation.

To avoid this difficulty about possible things, it may be held that all things which can be imagined are possible, though not necessarily conversely. Then, although we cannot conceive a particular object, we can very well imagine it, and the possible instances of <u>golden</u> <u>mountain</u> are all imaginable.

So the general term 'golden mountain' and 'square circle' are different in kind, the former being applicable to possible instances, while the latter being not applicable to anything at all.

But this way of interpreting the concept of possibility is not free from difficulties. For example, it is not clear if it will make sense to ask whether we can imagine the _same_ object twice--in the absence of any criterion of identity of imaginary objects. The imaginary objects cannot be identified with images without further ado, for it is clear that we cannot have the same image twice separated by a temporal gap, i.e., a criterion of identity of images is available.

Because of difficulties in distinguishing between particulars and universals by means of location and duration, an altogether different kind of distinction has been proposed by many. A particular, according to this theory, is a bundle or collection of universals, different particulars of the same type are different collections having one universal as a common item, and other universals in the collections varying from particular to particular. Thus there is no difference in kind in the stuff of which universals and particulars are made. This theory is sought to be justified on various grounds. It is pointed out that if we abstract the spatial and temporal aspects from a particular, we do not necessarily have a universal. F.H. Bradley, for example, held that the mere separation of the _that_ from the _what_ of a particular does not yield an ideal content which is a universal. The content thus separated from its existence has to be 'garbled', the special features of the content have to be rejected and only the common features retained, in order to yield the ideal content. That is, even without the _that_, we may have a particular content.

Husserl's epoché is a 'bracketing of existence' of particular objects. Yet this bracketing of existence by itself does not yield essences or universals--a further reduction, eidetic reduction, has to be performed to obtain essences. Thus a particular cannot be simply a concretisation of a universal in space and time.

This explains why in natural languages we do not have any method of forming abstract nouns from proper names. For the property of being Socrates, or Socratising, for example, is still a particular, though without existence and therefore may be used without any ontological commitment. But this does not make 'Socratise' a _general_ predicate, for it can be truly predicated of one and only one individual. So the phrase 'the thing that Socratises' contains an unnecessary 'the'.

A general term like 'man' denotes not merely those who actually exist at a particular time, but all men, past, present, and future. Past and future men may be conceived as possible men, i.e. as no longer, and not yet, actual; still it is not necessary to do so. Past and future men may be connected

causally, or otherwise objectively, with the actually existing men. Possible men cannot be causally related with actual men. The relation between actual and possible objects is not a causal relation.

Although all abstract properties are not universals, all universals in the sense of D1 are abstract properties, and may be denoted or referred to, by abstract terms formed from general terms. In every natural language there is at least one device of forming abstract nouns from concrete common nouns, although there may not be any uniform method of forming abstracts from complex concrete terms of arbitrary length. There is also a method of forming abstract nouns from suitable relational words, like the verb 'marry'. Thus we have the abstract noun 'marriage', and also abstract relational phrases like 'being the father of'. The question here is: Although such words and phrases are abstract terms, are they also used to denote universals? Strawson distinguishes between 'universals like marriage' and what he calls indifferently as 'universal-cum-particular' and as 'particular-cum-universal', like 'being married to John'.[7] Now Strawson regards expressions of this latter type as predicate expressions on the ground that such expressions 'as a whole, do not present any fact'. Yet it may be noted that 'being married to John' can be predicated truly only of at most one person (at a time) in monogamous societies (assuming that we refer to one individual by 'John'). That is, <u>being married to John</u> can be a property of at most one individual at a time, and hence is not a universal in the sense of D1.

Now what about <u>marriage</u>? Is it not a common property of all married individuals? It may be argued that the answer must be in the affirmative. For, from 'A is married to John' we can infer 'A is married' or 'A is married to some man'. Also from 'B is married to Smith' we can infer 'B is married to some man'; and from both together we may infer 'A has a property B has'. What is common to both A and B are that they are both married (to some man), which must, therefore, be a universal in the sense of D1.

But here we should note a peculiarity of such properties as <u>marriage</u>. Such properties are temporal properties which come into being at a particular time by virtue of one's performing certain formal rites. So the property of being married which A has cannot be numerically identical with the property of being married which B has, if John and Smith, (and/or A and B) are different persons.

Thus we have two different types of properties--properties which are timeless, and properties which are produced. Universals are timeless properties; humanity, for example, cannot be said to be produced; but marriages, as everyone knows, are <u>made</u> (even though in heaven). So we have 'marriages', but we cannot have 'humanities' (except in a different

sense). We can, of course, have a second order abstract term to denote what is common to all marriages ('marriage' being itself an abstract term of the first order), but what is common to all marriages will not be common to all married couples.

4. ANALYSIS AND UNIVERSALS

We now consider the problem of complex general terms of a simple type. Take, for example, the complex term 'blue lotus'. Just as 'blue' and 'lotus' are general terms, so is the complex term 'blue lotus'. The question which we want to ask here is: Is there a universal implied by 'blue lotus' which is different from the universals implied by 'blue' and 'lotus'? Occam's razor may be used here to avoid unnecessarily multiplying abstract entities. For a blue lotus may be regarded merely as something which is both blue and is a lotus. But this simplification involves difficulties if both blue and lotus are regarded as universals, of which the complex universal is composed. The difficulty here is about the nature of this process of composition, which may be brought out by contrasting two different types of complex terms of this sort.

First we have complex terms like 'rational animal' which have simpler forms usually of one word like 'man'. Then 'rational animal' is said to be obtained by analysing the simple term 'man'. The sentence 'man is rational' is called an 'analytic sentence.' Secondly, we have complex terms like 'blue lotus' which do not, as a rule, have simpler designations of single words. We may, of course, have a sentence like 'a blue lotus is blue' and call it 'analytic'. But it is important to note that the word 'analytic' is used in two different senses in the two types of cases. The sentences of the first type are analytic in Kant's sense, the predicate of which is 'covertly contained' in the subject, and a process of analysis is needed to reveal it. But in sentences of the second type the predicate is part of the subject most palpably, and no process of analysis is needed to bring it to the surface. So sentences of this type, if called analytic, will be analytic without needing any analysis. So complex terms like 'blue lotus' should not be regarded as identical in nature with those like 'rational animal'.

There is a more important difference between these two types of terms. In the case of 'rational animal', 'animal' is wider in denotation than 'rational', whereas this is not so in the case of 'blue lotus', for 'blue' is not wider (or narrower) in denotation than 'lotus'. The class of blue things and that of lotuses intersect or overlap. We shall say, in such cases, that the two universals themselves 'overlap'. The question here is: Can two universals overlap in this sense? The answer to this question depends on what function we want

universals to perform. If we want to use universals for classifying things according to their natural kinds, then we cannot permit universals to overlap; for overlapping universals will lead to cross-classification of objects and this is not permissible. But if we do not want to use universals for such classification, then there is no harm in admitting overlapping universals. It is usual here to classify universals (as common properties) into two kinds--(i) sortal universals, and (ii) characterising universals. Sortal universals are used to state what kind of thing a particular object is. Overlapping universals of this kind will put one and the same object into two different natural classes. But overlapping characterising universals do not have this consequence. There is no harm in asserting, for example, the following three sentences:

(a) Socrates is both ugly and wise.

(b) Plato is wise but not ugly.

(c) Smith is ugly, but not wise.

Here although ugly and wise are overlapping universals, yet as they are not performing the function of classification according to natural kinds there is no logical or ontological difficulty.

Sortal universals are often identified with essences. To say what sort of thing a particular object is, is to say at least a part of its essence. Characterising universals are not essences, they are not even properties (in the sense of traditional logic), for properties are like essences, and cannot be overlapping; they can only be accidents which happen to be common to many. A white man cannot be the meeting point of two essences, his color is not his essence or a property. The essence of white men is just the essence of men. The white color is only an accident and is common to all and only those things that are white, white men, snow, milk, etc. So there is no overlap of essences here, only an accident is overlapping with an essence.

If thus there cannot be overlapping essences, one essence can therefore be only wider or narrower than another essence, or completely disjoint from it. Given any two essences, either one will be completely subsumed under the other, or they will be mutually exclusive. There will also be no point in admitting two essences which have exactly the same instances; that is, there cannot be co-extensive essences. In such cases any one of the two will be regarded as the essence.

Now we examine the nature of the two processes of analysis and composition or synthesis which are inverse processes. A complex essence can be analysed into simpler essences, and two or more simpler essences can be synthesized into a more

complex essence. Both the processes, therefore, involve complex essences either as starting points or as results. Yet the nature of complex essences is not clear. The standard procedure of defining a term _per genus et differentiam_ involves the difficulty of understanding the meaning of _et_. How animality and rationality are inter-related to form the complex universal of humanity poses a problem. Baumgarten suggested 'transcendental' to characterise this relation. But his explanation of this technical term is not clear. There are difficulties in logically characterising it. For example, we may ask whether animality or rationality are accidentally or necessarily related in humanity. Their inter-relation cannot be accidental, for in that case, they will fail to form _one_ essence, humanity, but will be a loose and unstable complex. There cannot be any _reason_ for their inter-relation if it is accidental. If, on the other hand, their relation is conceived to be necessary, then there must be something in the very nature of essences (animality and rationality, for example), which make them necessarily united to form the new essence, humanity. The difficulty in this theory is to explain how two essences which are self-complete can also demand to be related to each other. This difficulty comes to the surface when attempting to explain how the same generic property can be necessarily related with different and mutually exclusive differentia to form mutually exclusively essences of coordinate species of the same genus. Discussing this problem, H.W.B. Joseph, for example, comes to this conclusion: "We may say that the genus and the differentia are one, because they were never really two.... The genus therefore could never exist independently of a differentia... nor the differentia of the genus."[8] But this deprives both animality and rationality of their essentiality, for no essence can be dependent for its existence on something other. Moreover, in thus emphasising the necessity of this inter-relation of genus and differentia, Joseph virtually denies that the genus is one essence. "So intimately one are the differentia and the genus that though we refer different species to the same genus, yet the genus is not quite the same in each; it is only by abstraction, by ignoring differences, that we can call it the same." Thus if we hold that animality and rationality are necessarily inter-related in humanity, we have different animalities for the different essences of cat, dog and the other coordinate species. Hence animality is not an essence.

Pursuing this line of argument one reaches the conclusion that essences cannot be complex, in the sense of being constituted by more than one essence. If there _are_ essences, then each of them must be one indivisible unity not merely in existence but also in knowledge. This leads to the position that essences can never be _analysed_, they are necessarily unanalysable simple entities; in other words, conceptual analysis of

essences into simpler essences is impossible. This applies only to essences or sortal universals, but not to characterising universals. There is no harm in admitting complex or composite characterising universals which can be analysed into simpler elements.

It is interesting to note that there is an aspect of Wittgenstein's theory which can be interpreted in support of this theory. According to him, what is common to all games is that they are games; no <u>further</u> analysis of any essence of games is possible. Champions of the conceptual analysis-as-the-method-of-philosophy find this position wholly unacceptable. A.J. Ayer, for example, complains, "It is correct, although not at all enlightening, to say that what games have in common is their being games." Enlightenment, if sought by further analysis, is not possible from the very nature of the case.

It is, of course, true that Wittgenstein intended his argument to be used not against the theory of analysability of essences, but against the postulation of essences, still this need not be the case. For, as Bambrough interprets Wittgenstein, he was arguing against the essentialists only because they meant by 'something common' 'something common <u>other than</u> being its instances' in answering the question: Must there be something common to all instances of a universal in order that they be denoted by one general term?[9] But this extravagance of the essentialists is not necessary. For, as we have stated before, the common property implied by a general term is referred to by the corresponding abstract term; and the abstract term corresponding to 'games' is only 'being games', to 'wise' 'wisdom' which is just 'being wise'. So where an abstact noun like 'wisdom' is available, 'being wise' is regarded only as a grammatical variant. Thus Strawson writes, "Let us say that the expression 'Socrates' serves to introduce the particular person, Socrates, into the remark, and the expression 'is wise' serves to introduce the quality wisdom, into the remark."[10] He also says, "It is different with 'is wise'. This expression introduces being wise just as 'Socrates' introduces Socrates."[11] Thus wisdom or being wise are the same, being introduced by the same expression 'is wise' into the remark. Now where an abstract noun corresponding to a general term is not available, as in the case of 'games', 'being games' is the only abstract term. Abstract nouns where available, mean only 'the property of being so-and-so or such-and-such'.

Although 'wisdom' and 'being wise' are both abstract terms formed from the general term 'wise', yet there is some difference between the two. 'Wisdom' denotes the property as abstracted from its relation to things which may possess it; 'being wise' on the other hand, includes the relation as abstracted from the things to which wisdom is related. That

is why we have 'Socrates being wise' as a proposition in Johnson's sense--an assertible containing the characterising tie, being; but 'Socrates wisdom' will not contain the characterising tie as 'wisdom' does not denote any tie or relation, but only the abstract property. But in spite of this difference, being wise and wisdom, as abstract properties, are analysable or unanalysable in the same way.

Thus there are abstract terms of three different kinds: (i) abstract terms, like 'marriage', which do not denote repeatable properties; (ii) abstract terms which denote analysable common properties like wisdom; and (iii) abstract terms which denote unanalysable common properties (essences) like humanity. It is only the second kind of abstract term which can be overlapping.

About the problem of the relation between universals and particulars it is to be noted that there are various aspects of this relation--ontological, logical, epistomological. As we have introduced the problem of universals as a problem of meaning of general concrete terms, our approach is primarily semantical and epistemological. We have also discussed the nature of abstract terms and their meaning, and the logical problem of complex essences and their analysability. We now conclude by remarking very briefly on the ontological aspect of the relation.

The question whether universals are in particulars arises only if both universals and particulars are regarded as real in the same sense. If particulars are regarded as shadows or copies or appearances, then their relation to the corresponding universals will be a relation between shadows and real things, between copies and originals, or between appearance and reality. Although these relations differ among themselves considerably, still it is clear that there cannot be any ontologically real relation between them, for a real relation can only hold between or among reals. So also if universals are regarded as concepts or ideas or as somehow being subjective, and particulars as real, then there also cannot be any real relation between them. If again universals are regarded as mere words, then the relation will be between arbitrary conventions behind the use of general words and the real world which is not arbitrary. In this case, too, it is clear that universals and particulars cannot have any real relation between them. All that our approach has shown is that universals being reasons for the application of general terms, must be epistemologically prior to particulars.

Calcutta University

NOTES

1. F.P. Ramsey: 1931, 'Universals', Foundations of Mathematics, Routledge, London, p. 113.

2. W.V. Quine: 1972, Methods of Logic, rev. ed., Holt, New York, p. 202.

3. Ibid., p. 207.

4. Ibid., p. 207.

5. Ibid., p. 130.

6. Ibid., p. 131.

7. P.F. Strawson: 1959, Individuals, Methuen, London, pp. 137, 189, 175.

8. H.W.B. Joseph: 1916, An Introduction to Logic, 2nd ed., Clarendon, Oxford, p. 83.

9. R. Bambrough: 1966, 'Universals and Family Resemblances', in G. Pitcher (ed.), Wittgenstein, Anchor Books, Garden City, N.J., p. 198.

10. P.F. Strawson, Individuals, p. 146.

11. Ibid., p. 149.

J.N. Mohanty

PSYCHOLOGISM IN INDIAN LOGICAL THEORY

The Indian logical theories, the Nyāya in particular, offer interesting models in the light of which one may hopefully throw new light on some of the persistent problems in philosophy of logic, which are generally discussed within the framework of Western thinking. The new possibilities that they open up could be instructive, at least by way of extending the boundaries of our available models, and we may be able, as a consequence, to see some of the limitations under which discussions in philosophy of logic are carried out. I want to discuss here one such problem: the issue about psychologism in logic.

1.

To begin with the issue of psychologism.[1] How is logic related to psychology? Briefly stated, psychologism is the view that theoretical foundations of logic lie in psychology. If one appeals to the rather commonplace distinction between how one ought to think and how in fact one does think, psychologism insists that ideal, logical thinking is but a species of thinking in general, and so is governed by the same rules that hold good of all thinking. The anti-psychologist philosophers, such as Frege and Husserl, argue, as against psychologism, that it in effect reduces the necessary truths of logic to the inductive, probabilistic laws of psychology; that it confuses between the laws of being-true and the laws that determine taking-something-to-be-true, between the objective entities with which logic is concerned (such as concepts, propositions, theorems, theories) and the subjective and private events in peoples' minds, and finally that psychology as a science, i.e., as a theory, presupposes logic rather than vice versa. While psychologism is clearly at fault in seeking to derive the laws of logic from the way the human mind works (note that the idea of 'deriving' itself involves the logical principles), the extreme anti-psychologistic position leaves us with an unmitigated Platonism in the strong sense, a domain of abstract entities such as meanings, propositions and theories, sundered, on the one hand, from the mental acts of thinking and understanding which grasp them and, on the other, from the linguistic signs which "express" them. Is there any way to avoid these two alternatives?

B. K. Matilal and J. L. Shaw (eds.), Analytical Philosophy in Comparative Perspective, 203–211.

The anti-psychologistic philosopher such as Frege is right in his intuition that the entities with which logic deals could not be privately owned, temporally individuated particulars, and that the logical truths are not inductive generalisations. But, it may be pointed out, he has an impoverished conception of mental life. This is certainly true of Frege, for whom the mental is the private particular and the laws about the mental are necessarily inductive generalisations. The question then for us is: why need we restrict ourselves to this impoverished concept of the mental? Is it not possible, given a suitably enriched philosophy of the mind, to give psychologism its due, while preserving the intuitions on which the anti-psychologistic positons are based?

I think that is possible. Indeed, there are several such systems already available. One is the Brentano-Husserl conception of an eidetic psychology. Another is to be found in the Indian logical theories, especially the Nyāya. In any case, what one needs is a conception of the mental according to which a mental event (or act) exemplifies or embodies a universal structure such that two or more numerically distinct mental events may exemplify or embody identically the same structure. In that case, within a mental event, one may distinguish between its particularity (which it shares with none other) and its universal features. Amongst universal features are to be counted not only such generic features as the property of being a belief or the property of being a desire, but also such specific features as 'being a belief that p' or 'being a desire for a cup of tea'. In brief, what one needs is the conception of a structure of mental acts. Once you have such structures, then one can talk about essential truths about different sorts of mental acts. In that case, a logic of propositions and an eidetic psychology of mental acts in which those propositions are entertained would have a closer relation than the radical anti-psychologism of Frege would like to admit, and yet both that relation and the nature of the relata would preclude one from falling into the obvious errors of psychologism.

In the light of these general remarks, let me now give a brief sketch of the Nyāya concept of (mental)[2] acts. If m is a mental act, it has an owner, i.e. a self, it occurs at a time, it has what Husserl calls an act quality[3] (i.e. it is either a perceptual act or an act of remembering, or an inference, or a desire, or a hope, and so on) and finally it has a structured content (about which I have more to say below). Of course, the act has an object, but the object falls outside the act. What represents the object within the act is the structure. The act m may then be represented as an ordered quadruple [self, t, q, content], where 't' stands for the time of the occurrence and 'q' for the act-quality. Our present concern is with the content.

The content of an act is neither the object of that act (which is, on Nyāya realism, always out there in the world) nor a real constituent of the act (which, in Nyāya ontology, has no parts, and so is formless, nirākāra, if 'form' signifies the structural arrangement of parts). To begin with, let us call it--using a concept handily available in phenomenology--'intentional content.' It is the object not as it is in itself but precisely as it is being presented in the act under consideration. The object out there in the world may remain the same, but for a different act having the same object, the structure of the content may very well change, depending on how precisely that object is presented in the new act. Consider the two perceptual acts expressed in the sentences:

(1) '(This is a) blue jar.' ('nîlo ghatah')

(2) 'The blue (is) of the jar.' ('ghatasya nîlah')

It would appear that whatever is the object out there is not changed by the change of one's perspective. In (1), the primary object is the jar which is being perceived as qualified by the colour blue. In (2), the primary object is the colour blue which is being perceived as belonging to the jar. I will not expound in detail the Nyāya analysis of these structures.[4] It would suffice to note that for the Nyāya such a structure is a concatenation of a whole set of (epistemic) entities, each of which serves a specific function of either qualifying/determining/limiting or being qualified/determined/limited by some other. These peculiar entities are epistemic, for they do not exist in the object per se, they arise only when a cognitive act is directed toward the object. These entities as well as their concatenated structures are, in an important sense, universal-like: another cognitive act may embody precisely the same structure. What the logician has directly to deal with is a cognitive act insofar as it exhibits such a structure. We have then a criterion of identity for acts for the purposes of logical analysis. Two acts m_1 and m_2 are l-identical, if they have the same act-quality and the same structure of their contents. The fact that they occur at different times and/or in different selves, is irrelevant.

At this point, one may wish to argue that the supposedly repeatable structure is nothing other than the proposition of Western logic. Two numerically distinct acts of belief are then l-identical in so far as they are beliefs in the same proposition. If the structure is nothing but the proposition, and since the proposition is an abstract entity towards which one may take different attitudes, or the same attitude at different times--then a logic of propositions would have to be

separated from a psychology of those attitudes. Now to appreciate the nature of the Indian logical theories, it is important to see why the content of a mental act as understood in the Indian logics is not the proposition of Western logic, not at least in one of the senses of 'proposition.' In this sense, which is also the sense in which detaching the proposition from the mental acts may be most persuasively effected, a proposition is that entity towards which many different, numerically as well as qualitatively different, attitudes and acts, belonging to the same or to different selves, may be directed. Now on the Nyāya analysis of the content of an act, the quality of the act does often make a difference to the content. In the sense of 'proposition' just indicated, the supposition 'S may be P', the question 'Is S P?' the denial 'S is not P' and the affirmation 'S is P' are all directed towards the same proposition. This is not the case in Indian logic, where analysis reveals a different structure in the case of 'Is S P?' than in the case of 'S is not P', and a different structure in the latter than in 'S is P'.[5] But the affirmative categorical 'This mountain has fire' does express the same content, not only when it expresses the cognitions of two different persons, or of the same person at different times, but also when it expresses cognitions of different types: perception, inference, or _śabda_. This justifies bringing these under one generic group called '_anubhava_.' This is not to deny that there are attempts to still more finely individuate the content even across these variations, so that the structure of the content would be different in the case of a _śābda_ knowledge from that in the case of an inferential knowledge, both again different from the structure of a perceptual knowledge. The 'proposition' of Western logic is not as finely individuated across the range of varying propositional attitudes.

There is still another difference between 'proposition' and the 'content' of Indian logic. Proposition is an abstract entity towards which a mental act is directed. Irrespective of how strongly one may want to ascribe to it an ontological status, it is independent of, and transcends that, or in fact any act directed towards it. But the content which one, through reflective analysis, discovers in an act, is that act's structure, not its object, not a transcendent entity.

Let us now see how this applies to the case of an inferential knowledge with which logic is concerned in the first place. This would involve determining in what sense the theory of inference proposed is, or is not, psychological. Consider the following account to be found in the Nyāya treatises on inference:

One sees smoke on a distant mountain. This leads him to remember the rule "Wherever there is smoke, there is fire" which he recollects as having been instantiated in cases such

as the familiar stove in the kitchen. He now recognises the smoke he saw on the mountain as a mark of fire in accordance with the rule he just remembered. At this point, if there is no unexpected hindrance, the person would, as a matter of course, be led to draw the conclusion: "therefore, there is fire on this mountain." This last sentence is an expression of an inferential cognition that has been produced in him.

What we have is a sequence of psychological events: a perception, a remembrance, a recognition, leading finally to an inferential cognition. These events belong to one and the same self, and are individuated both by ownership and temporal position. Now how can any such temporal sequence yield a logical rule? We can do that by (i) replacing the particular person concerned by a variable and making a universal quantification over it; (ii) by retaining appropriate relations of succession, but doing away with the actual temporal positions; (iii) by identifying the cognitions involved by their contents and temporal positions relative to the other cognitions figuring in the rule; and (iv) by requiring that all cognitions figuring in the rule must have one and the same owner. We then get a rule such as the following:

> (3) For any knower *S*, if *S* has a perceptual cognition *Fx*, and then remembers the rule, "Wherever *F*, there *G*: as instantiated in the uncontroversial case *O*", and then perceives in *x* the same *F* as before but this time as figuring in the remembered rule "Wherever *F*, there *G*", then *S* will experience an inferential cognition of the form *Gx*, *provided* there is no relevant hindrance.

This indeed is as much a law in eidetic psychology as one of epistemic logic of inference. It is arrived at by an intuitive induction over particular cases, it is not a probabilistic inductive generalisation.

Another set of such laws with which the Nyāya logic operates consists of the rules of the form:

> (4) If a cognition of the type *Ø* and with a structure *T* occurs at time t_n in a self *S*, then a cognition of the type ψ and with a structure *T'* would not occur at time t_{n+1} in *S*.

A simple and intuitively clear case of such a law is: if a person perceptually ascertains that -p, he cannot, at the immediately following moment, have a perceptual cognition that p. Or: if a person perceptually ascertains that p, that perceptual cognition will prevent the emergence, at the immediately succeeding moment, of an inferential cognition that -p

(even if other conditions for the latter cognition are present).

These rules are further strengthened by bringing into consideration non-cognitive causal conditions such as desire to have a certain sort of knowledge. One cognitive type θ is said to be stronger than another cognitive type ψ, just in case if the causal conditions of both are present the one belonging to the type θ will occur and the one belonging to the type ψ will be prevented from occurring. Thus, if the causal conditions for a perceptual cognition of the fire on yonder mountain are present, as well as conditions for an inferential cognition of the same in the same person, then the perception will occur, preventing the inference from occurring. But suppose, in addition, that there is a desire to infer: presence of this new factor will cause the inferential cognition to occur even if the perception has just occurred. One may infer "There is fire on the mountain" even if one just saw the fire on the mountain top, if only one desired to so infer.[6]

Given such an eidetic psychology of cognitions of which rules such as the above are fundamental principles, one can have a theory of inference which is indifferently a logic and a psychology of inference, but which is not 'psychologistic' in the pejorative sense. But given this formulation of theory of inference, a serious question arises. If the rule formulated in (3) is _also_ a psychological law, then it would seem as though all persons would necessarily make the right sorts of inferences under right conditions, and it would be nearly psychologically impossible to commit a logical fallacy. For if the logical rules of inference are also rules of appropriate cognitive occurrences, then it would be impossible for men, given the psychological constitution that we have, to violate those rules. Now this indeed is the most difficult question for a psychologistic theory of logic to answer. However, let us try. The place we need to look at is the theory of invalid inference, or what has been called _hetvābhāsa_ (to be construed as either defective _hetu_ or the defects of _hetu_, _hetu_ being the mark from which an inference is made). The standard definition of a defective _hetu_ which would vitiate the inference in which it functions as _hetu_ and would render it invalid is this: "the object of such knowledge as acts as the preventer of inferential cognition."[7] What is intended is, in brief, this: an inferential cognition of the form "S is P, because of m" would be prevented from occurring if the person under consideration has a knowledge of a situation which is, in fact, a defect of m _as a mark_ of P-ness in S. Consider the obviously fallacious inferential cognition, "This lake possesses fire, because it possesses water." Such an inferential cognition can occur in a person only if he or she believes in the truth of the universal rule "Wherever there is water,

there is fire." However, if the person recognises that wherever there is water there is the absence of fire, which amounts to recognising that the hetu or mark is characterised by the defect known as 'viruddha', or that the hetu is a viruddha hetu, then the inferential cognition would be prevented from taking place. This is rather a curious way of putting the matter. Instead of being told that the person made an inference that is fallacious, we are rather told that he or she would not have made the inference if only he or she had known that the hetu that was being employed was defective. One way of understanding all this--the one I prefer, for it meshes well with the account I have developed earlier in this paper--is to take the thesis to imply that as rational beings we cannot make a fallacious inference, we only appear to be doing so. Since the causal conditions of inference require, in accordance with (3), that the person concerned must believe in the appropriate rule "Wherever there is m, there is p", he or she can infer only if there is such a cognitive occurrence in his or her mind, so the inference he or she makes will always be formally valid. Now that he learns that in fact "Wherever there is m, there is -p", this cognition will prevent that other rule-cognition and so eventually the inferential cognition from occurring. The implication of course is clear: even when we are apparently making an invalid inference, we are making it because we not only do not detect the fallacy involved but also because we are so construing the terms and the premises involved that the inference would turn out to be valid. Since psychologically it is impossible to make a fallacious inference, when we make an inference which by objective criteria is fallacious, what is happening is that we have given the premises and the terms, interpretations under which the logical and the psychological requirements are in fact satisfied. If those interpretations are changed--and this is what happens when the defect in the mark is recognised or pointed out--that inferential cognition would be prevented from recurring.

This is the price one pays for making the psychology of cognitions and the logic of propositions to coincide at least within the limits of elementary inferential operations. There is a concomitant commitment to rationality which rules out the possibility of making such obviously invalid inferences as "All men are mortal, Socrates is a man, therefore, Socrates is not mortal." However one who does make such an inference must be misconstruing the senses of the logical terms "All" and "not".

The thesis is not as improbable as it may look to be at first. Mary Henle has found out, by considering empirical data about errors in syllogistic reasoning by adults, that "where error occurs, it need not involve faulty reasoning, but may be a function of the individual's understanding of the

task or the materials presented to him."[8] In another experiment, this time with children, Henle fails to find evidence that thinking transgresses the rules of syllogism.[9] In most cases, the subjects understood the premises in a manner that accounted for the error, while no faulty reasoning process was employed. The implications of her findings, as Henle sees them, are that "the two blind alleys of psychologism and of the radical separation of logic from the study of thinking" have to be avoided. Saying that our actual thinking process exhibits an (implicit) logical structure does not, in her view, amount to psychologism, for it does not make "logic coextensive with thinking by making it illogical. Rather than denying logical requiredness, denying the demands of necessary implication, it seeks to show that such requiredness is central in actual human thinking."[10] Such a conception of actual human thinking, I want to emphasise, is germane to the Indian logical theories, especially the Nyāya which finds the logical in the texture of everyday actual processes of reasoning. This is done, as we have seen, by construing the mental processes of reasoning as rule-governed patterns of succession of cognitive events (jñānāni), the rules being not empirical generalisations but Brentano-like intuitive inductions.[11]

I would like to add, at the end, that the logical structure of a cognition should not be taken to coincide with the structure of the sentence which expresses that cognition, for one reason amongst others that there always shall be constituents of the cognition--e.g., the mode of presentation (a la Sibajiban Bhattacharrya = Fregean Sinn)[12] --which cannot be expressed but can only be shown in that sentence. In other words, for an expressed sentential constituent, there necessarily shall be an unexpressed epistemic constituent. This should not be construed as suggesting an ineffability thesis, for what is unexpressed in that sentence can be expressed in another which, on its part, shall have its own unexpressed epistemic content. A given sentential structure does not then provide a clue to eliciting the epistemic structure unless it is aided or rather supplemented by reflective analysis of one's own cognition. Structural analysis and reflection on the inner cognitive events are, ideally, made to supplement each other.

Here we have a possibility which neither Frege nor Husserl, in their eagerness to reject and overcome psychologism, saw; but Husserl was closer to seeing it than Frege.[13]

NOTES

1. 'Psychologism' here should be understood in that sense or cluster of senses in which Frege and Husserl used it.

2. The acts are not, strictly speaking, mental for the Nyāya; for in the Nyāya ontology mind (manas) acts as an instrumental cause in the production of the acts, while the acts belong not to manas but to the self. Extensionally, though, we may identify the acts as 'mental acts'.

3. Cp. E. Husserl, Logical Investigations, vol. II, Investigation 5. E. tr. J.N. Findlay.

4. Cp. J.N. Mohanty: 1966, Gangeśa's Theory of Truth, Santiniketan, Introduction; and B.K. Matilal: 1968, The Navya-Nyāya Doctrine of Negation, Harvard University Press, Cambridge, Mass.

5. For more on this, cp. Gangeśa's Theory of Truth, Introduction.

6. For more on such rules, cp. Sibajiban Bhattacharya, "Some Principles and Concepts of Navya-Nyāya Logic and Ontology," Our Heritage 24.1, 1-16 and 25.1, 17-56.

7. "Anumitipratibandhakayathārthajñānaviṣayatvam hetvābhāsatvam" Annambhatta in his Dîpikā on Tarkasamgraha.

8. M. Henle: 1962, 'On the Relation between Logic and Thinking', Psychological Review 69, 366-378, especially p. 373.

9. M. Henle: 1971, 'Of the Scholler of Nature', Social Research 38, 93-107.

10. Ibid., p. 107.

11. For Brentano's concept of eidetic psychology, see R.M. Chisholm: "Brentano's Descriptive Psychology," in The Philosophy of Brentano, ed. Linda McAllister, Humanties Press, Atlantic Highland, New Jersey, 1977, pp. 91-100.

12. Cp. Sibajiban Bhattacharya, loc. cit.

13. Cp. J.N. Mohanty: forthcoming, 'Husserl and Frege on the Overcoming of Psychologism'.

Karl H. Potter

A SPEECH-ACT MODEL FOR UNDERSTANDING NAYVA-NYĀYA EPISTEMOLOGY

1.

Is there philosophical analysis in Indian thought? The question is interpretable in numerous ways, among which are these: (1) do Indian philosophers attempt to clarify concepts by analyzing them into their components? and (2) does one find in Indian philosophy concepts of the sort used by analytic philosophers in the West? I'm quite sure of the answer to (1); it is in the affirmative. One can easily locate attempts to clarify through analysis in Indian thought. Unless one places further restrictions on what is going to count as "analysis", one can find such attempts in just about anybody's thought: it would be surprising if any sophisticated culture was so uninquisitive that its members didn't regularly ask each other what they were talking about.

So I shall assume that any interesting formulation of the question of philosophical analysis in India must turn on some more specific notion of what an analysis involves. But that leads directly to the following question: is "philosophical analysis" a culture-free term standing for any culture's ways of going about analyzing concepts, or is it culture-bound in that it refers to the categories in which those people traditionally called the analytical philosophers went about analyzing things? In the culture-free sense India has its own methods of philosophical analysis--indeed, perhaps a host of different methods corresponding to different persuasions of the different systems--and to show instances of it (or them) one should select passages and sketch the analysis provided. If one is, let us say, a member of a philosophical school pioneering or explicating such an analysis for the benefit of other members of the school, one will address this question using the linguistic and conceptual tools that his audience already understands. The question of whether these tools correspond to or contrast with the tools other systems, or other cultures, use just doesn't arise unless someone goes out of his way to ask it.

But suppose someone, let us say a member of the Nyāya-Vaiśeṣika tradition in India, is asked to explain his tools of philosophical analysis to someone outside of his, or indeed of Indian, culture. He may, of course, merely say "you'll

B. K. Matilal and J. L. Shaw (eds.), Analytical Philosophy in Comparative Perspective, 213–230.

have to learn Sanskrit; these tools that I use are unintelligible except as expressed in their indigenous terminology, which is Sanskrit." There is a lot to be said for this reply, and I suspect it is ultimately the only satisfactory reply if the expected explanation is to evoke the deep understanding which will, e.g., allow the audience to instigate fresh analyses of their own without regularly committing howlers along the way. That is to say, I think that comparisons with other analytic tools whose home is in another place can never fully replace the indigenous understanding of those tools by someone of the home territory.

Still, in communicating about one's indigenous analytic tools there is clearly a place for drawing comparisons, and they are extremely useful, possibly indispensable, in getting the audience to think along lines with which they are familiar and which show the strongest analogies to the way of thinking which one is trying to explain. So it is of some importance, in explaining another culture or system of philosophical analysis to an alien audience, to pick as appropriate a comparative model as one can.

It seems fair to say that analytic philosophy had its inception in the methods of Frege, Moore, Russell, various logical positivists and others of that persuasion. There is quite a growing literature involving attempts by scholars of Navya-Nyāya to explain the concepts used in that school to an English-language audience utilizing comparisons with such analytic philosophers as those just mentioned. For example, a number of us have mentioned an alleged agreement between the spirit of Russell's theory of descriptions and the implications of Nyāya analysis.[1] Books have been written comparing Indian theories on "logical" matters with those of writers such as Russell, Quine, Wittgenstein, and others.[2] Confronted with the evidence of this literature, one might begin to hazard the guess that the analytic philosophy practised by "rational reconstructionists" provides a case for thinking that there is a real, rather than merely a heuristic, ground for comparing Nyāya with analytic philosophy, a ground consisting in actual similarities.

The reasons why one picks a set of concepts, or an analytical method, as a model on which to base an explanation vary from case to case. I suspect it mainly has to do with familiarity, a familiarity which frequently turns on one's academic training. Someone trained in existentialist thinking will find analogies between Indian thought and existentialism; Thomists will cite Aquinas; and so on. Analytic philosophy has no more claim to be a natural model for Indian thought than any other, and as a heuristic mechanism for easing the pain of attempting to understand an alien conceptual system its usefulness is largely dependent on its familiarity. Yet I do believe there is a sense in which one can ask, as among

different models drawn from an alien culture, which of them fits the system to be explained best, or at least as between two such, which fits better. In this paper I want to explore the relevance to understanding of Nyāya of two models.

The two models may or may not be thought of as branches of analytic philosophy. I suspect they are usually so considered, but there are those who view the speech-act thought of J.L. Austin as part of a reaction to analytic philosophy rather than as an extension of it. In any case, I am interested in contrasting classical analysis of the sort practised by Gottlob Frege, Bertrand Russell, Rudolf Carnap, W.V.O. Quine and those in general who talk about "propositions", with Austin's style of analysis which places emphasis on the activities we perform with words, and to argue that when the most fundamental concepts for Navya-Nyāya epistemology are studied it is the latter, rather than the former, style which provides the closer and more instructive basis for comparison.

It must be kept in mind throughout that we are exploring what are at best analogies here. As was said, there is an important sense in which a concept is inextricably tied to the language in which it is expressed, so that a term whose home is in Sanskrit can in no way be equated with one whose home is in English. Rather, one must first understand each concept in its home language; only then can the sort of project undertaken here be a responsible one. So I shall perforce consider the uses of the relevant Sanskrit terms in the course of trying to gauge whether one or the other model fits best in explaining the thought in which that term figures.

2.

The basic term for our purposes is the term jñāna. A number of writers have pointed out the problems in translating this term as it occurs in Nyāya. It is regularly rendered as "knowledge" (a mass term), but what is shown below will, I think, clearly demonstrate how misleading a translation that is. I would suggest (the count term) "awareness" as the best English translation I have been able to come up with.

Now in Nyāya it is jñānas which are our thoughts, ideas, our epistemic states. A jñāna is not an abstract entity like a proposition; it is the momentary psychological state we are in when we are experiencing either a perceptual or a conceptual awareness. More than that, a jñāna is an act. It has an agent, a purpose, a result. In these respects it is like a speech-act and unlike the state of entertaining a proposition.

In classifying types of actions (karman) Indian tradition has regularly alluded to three types: bodily (kāyika), mental (mānasika), and vocal (vācika). The Nyāya theory is that a jñāna is a mental act. In fact, the same view is accepted

by just about every kind of Indian philosophy.

The sense of "act" in which a jñāna is a mental act is analogous to that in which a speech-act is an act--it involves an intention on the part of its agent. This intention need not be recognized as such by the agent at the time of the act, nor indeed ever, and no overt effort need be made or experienced. Nevertheless, there is activity by the agent of a mental act, just as there is by the speaker of a speech-act, in that they have attended and responded in a specifiable way on some occasion.

In Nyāya, jñānas range from minimal perceptual awarenesses of the sort termed nirvikalpaka (which awarenesses we are never aware of except by inference after the fact), through ordinary (savikalpaka) perceptual awarenesses, to what are termed awarenesses expressed in language (śābdabodha), inferences, and on to more complex conceptual awarenesses. The notion is that we initially grasp the elements of what we are aware of in what we may call "atomic" (nirvikalpaka) fashion: e.g., we may be aware of a particular somethng, of a universal property, of a certain quality, or a specific relation, but in each case independently of anything else. The simplest awareness that we can be directly (non-inferentially) aware of is already "molecular" (savikalpaka), involving at least a something qualified (viśesya), a qualifier (viśesana) and a relation between them (sambandha). The Naiyāyika hypothesizes that there must have been the previous, "atomic" stage, since otherwise how could we have gotten the qualificand, the qualifier and the relation into our consciousness?

All awareness, whether atomic, molecular, linguistically expressed or not, have contents (visaya). As we have just seen, the content of a nirvikalpaka jñāna is an unrelated item, the content of a savikalpaka jñāna is at least an interrelated triad. What is "content"? For Nyāya, it is an actual entity, a feature of the real world. Nyāya's is not a dualistic epistemology; it does not countenance something, an "idea", which mediates between awareness and the world. The arguments for dualism stemming from the fact of perceptual error are known to the Naiyāyikas, but they reject them. As they see it, the mind has no power to project contents--even ideas--into the world. Contents are always exhaustively composed of real items. Even in illusions and hallucinations what is presented to awareness is real: these experiences are erroneous not because they apprehend unreal contents, but because they misrelate spatially and/or temporally the objects they apprehend.

A jñāna, then, is an act of awareness intentionally directed at one or more actual entities. It is called "true" (pramā) just if the entities in question are related spatially and temporally the way the awareness takes them to be.

If that is not the case, the jñāna is termed "not true" (apramā). One should not assume that only erroneous awarenesses are apramā; that category also includes certain other cognitive states which are not true in rather different ways. For example, doubt (e.g., "is it a man or a post?") is apramā, since it is not the case that a doubt is a true awareness in the sense mentioned, but we should not call a doubt an error.

These, then, are some of the features of cognitive awareness as conceived by the Naiyāyikas. My contention is that they match the implications of a speech-act epistemology better than one of the classical, propositional variety. Let me now discuss some of the contrasts and comparisons which lead me to put forward this contention.

3.

Here are some features of the Nyāya conception which seem to me to suggest parallels with speech-act theory.

3.1. A jñāna is an act, as we saw: it has an agent, a purpose, a result. Likewise, a speech-act has an agent, a purpose, a result.

3.2. Not all jñānas are beliefs. Some are. In Nyāya one can pick out those jñānas which are beliefs by saying that they involve niścaya or vyavasāya, terms meaning "ascertainment"--beliefs are awarenesses in which the agent makes a claim to understanding, as opposed, e.g., to doubting (saṃśaya), where no such claim is made. The parallel in speech-act theory is with asserting, which is a kind of speech-act in which the speaker makes a claim on his hearer's beliefs.

3.3. A jñāna has a content. A speech-act has intentionality: it is directed toward something. Different versions of speech-act theory have different terminology for this something: Hare calls it the "phrastic"[3] (earlier he called it the "descriptor"[4]), Searle calls it the "proposition".[5] In either case this content is what two speech-acts which are tokens of different kinds of act may have in common as a type in virtue of what they represent. E.g., "Jones! Shut the door" is a command, while "Jones will shut the door" is an assertion, but they have the same kind of content, the state of affairs of Jones' shutting the door in the future. Likewise, my thought at t_1 that Felix is on a mat at t_1, and your thought at t_2 that Felix is on that mat at t_1, have the same kind of content, the state of affairs that Felix is on a certain mat at t_1.

3.4. The meaning of a jñāna is a function of pragmatic considerations which determined relevant semantic and syntactic features. The Nyāya theory of the meaning of a verbal awareness (śābdabodha) attests to this. According to that

theory, meaningfulness is a function of four considerations: expectancy (ākāmksā), fitness (yogyatā), contiguity (samnidhi) and intention (tātparya).[6] Expectancy is a matter of syntactic or "logical" well-formedness: the items constituting the content of a verbal awareness must involve a qualifier and a qualificand, as well as a relation to link them, or else one has not got a verbal awareness at all, but some other kind of awareness, perhaps a confused one. The parallel to this in speech-act theory is that a nonsense string of syntactically unconnected words will (except in unusual situations) perform no communicative function and so be without meaning. Fitness is semantic well-formedness: the qualifiers and qualificands and relations must be of the appropriate sort to perform their function, or else the jñāna will be without meaning since it fails to function in the way intended by its agent, whatever way that is (whether as ascertainment, doubt, or whatever). Contiguity has to do with the connectedness of the elements in the jñāna; if the elements are spatially or temporally far removed from each other, again the jñāna cannot fulfill its function(s). Finally, in cases where the above considerations are satisfied but the practical functioning of the jñāna is blocked by ambiguities in the modes of expressing it, appeal is made to the agent's intention as determinative of the meaning of the jñāna, i.e., as to which jñāna he has in mind. The entire picture is one of the agent's performing an act of awareness involving that act's being of a certain kind and having a certain content; the success of the act is contingent on the act's satisfying the four conditions mentioned, that success consisting in the agent's making a judgment of one kind or another about something and that judgment's being assessable in the relevant way. If the judgment is one of ascertainment (a belief) the assessment in question relates to its truth or falsehood, if it is a doubting, the assessment will be with regard to whether it should be resolved by a belief in one or the other of the options (a man, not a post).

3.5. Truth is a property of the jñāna, not of its content. The terms in the content of a jñāna, as we have seen, are for Nyāya actual entities. No question about their truth can arise; they just are, exist. Truth is a matter of a certain fit between the jñāna and its content, a fit which as we have seen requires that the jñāna not only have a content but that it be a jñāna of a certain sort, viz., a belief. A true jñāna is one in which the actual entities are represented as related spatially and temporally in the manner they actually are related. A false belief is one in which the same kinds of actual entities are represented as spatially and temporally related in a manner other than the way they actually are related.[7] This conception is at least partially paralleled in speech-act analyses such as Austin's, where it is the

entire speech-act which is the vehicle of truth. I cannot, of course, say that speech-act analysis as it has been developed by Austin and others has been characterized by the strong realist epistemological assumptions which Nyāya makes, but it is at least open to such an interpretation, whereas, as I shall go on to discuss shortly, classical analytic theory makes this realistic epistemology virtually impossible.

3.6. A jñāna is a mental act with a specific viṣaya or content. It has that content as a part of its very nature, whether it is (properly) understood to do so or not. That is to say, the content of a particular mental act is fixed, not by how it is described or is thought about by people (even including the agent himself) but in virtue of which mental act that jñāna is. Likewise, a particular speech-act may be conceived as individuated by its phrastic and neustic together, that is, both by the kind of speech-act it is (asserting, doubting, commanding, etc., which Hare calls its "neustic"[8]) and by the particular content or phrastic it has in virtue of its speaker's intention, whether the speaker (or anyone else) describes the act in appropriate terms or not. Once again, it is true that not all speech-act analyses have insisted on this point,[9] unfortunately to my mind, but it is a possible line to follow in speech-act analysis, whereas in classical analysis it is not.

3.7. It follows that the intention of the agent of a jñāna is determinative of the meaning of the jñāna, just as the speaker's intention is determinative of the phrastic of his speech-act (according to those types of speech-act analysis which makes the speaker's intention fundamental). So it is possible for a speaker or agent of a jñāna to say correctly "I thought I had believed (meant) p, but I was wrong; it was really q", whereas on the contrary account the belief (utterance) p must be (mean) something different from q, making the quoted remark above literally nonsensical. Nyāya analytical terminology clearly assumes the interpretation suggested here when it speaks of the "limitor" (avacchedaka) of the contentness (viṣayatā) of a jñāna. A content may be viewed in its own nature, or "under another description", as one would put it nowadays. The Nyāya way of making this distinction is to say that in the former cases the contentness is limited by its own nature (or "by itself"), in the latter it is limited by a property other than its own nature. The old Latin particle "qua" was used to the same effect.

3.8. A jñāna may be considered in the light of its several roles (1) as a quality (guṇa) of the self or agent; (2) as an instance of a certain type of awareness and a certain type of content; (3) as productive of certain results, intended or unintended. These roles nicely parallel the tripartite analysis made famous in Austin's writings between the locutionary, illocutionary and perlocutionary dimensions of a

speech-act.

4.

In classical philosophical analysis of the sort pioneered by Frege, Russell, Wittgenstein in his early days, and the positivists, an awareness such as a belief was understood as a relation between a person and a timeless entity called a "proposition". Thus "A believes p" made belief a relation between A and p. A proposition, on this reading, is not the content of a particular belief by nature; it is identified and individuated in a different way. The notion of a proposition is introduced through consideration of the point that it is possible to say the same thing by using different words or words in different orders. The proposition is the thing that can be expressed in different ways, what is expressed by, say, both "Calcutta is east of Delhi" and "Delhi is west of Calcutta". Alternatively put, it is that information which is expressed in all statements having the same logical entailments.

Thus a proposition, for the classical theory, is basically something abstracted from language, and specifically from the language of assertions. It has a life of its own, independent of anyone's asserting it or believing it. Whereas Nyāya views cognitive judgments as analyzable into a self related to a jñāna (at time t) by the relation of inherence (samavāya), classical analysis views it as a person related (at t) to a proposition by the relation of believing. As a result, one can study the relations among propositions independently of the epistemic contexts in which they occur, and such a study, as construed by these classical analysts, is logic.

4.1. The logic thus construed is a logic thoroughly abstracted and separated from thought as a process. A proposition is defined in terms of its logical entailments. Thus if A believes proposition p he *ipso facto* believes all those propositions entailed by p. In this dispositional sense of "believes" I believe a lot of things I'm not aware of believing, that I do not believe in the episodic sense of undertaking an occurrent act intentionally directed toward them. Believing is not an act for classical analysis; it does not have an agent, a purpose or a result. Episodic believing is a matter for psychology, not for philosophy.

4.2. Let us then distinguish between (episodic) believing-e as an act, and (dispositional) believing-d as a relation between a person and a proposition. For classical theory, all acts of awareness of whatever sort involve believing-d, and so classical analysts confined their attention to beliefs-d, viewing the resulting analysis as fundamental to all epistemic attitudes of whatever sort. In Nyāya the attention is to

believing-e which results in attention in turn being paid to epistemic attitudes of other sorts in their distinctive natures. In parallel fashion, in speech-act analysis asserting is an occurrent act, one among many neustical varieties, not a standing condition of any linguistic act, and so likewise one finds there attention paid to the varieties of linguistic acts.

4.3. The classical theory of a proposition does not think of propositions as the contents of beliefs-e in the way that speech-act theory views phrastics. Jones' shutting the door in the future is not a proposition for classical analysis. "Jones! Shut the door" does not entertain a proposition at all on that theory. The proposition rather is something which has the form of an assertion (say, "Jones is shutting the door") but is not itself linguistic. Furthermore, whereas it makes a certain kind of sense to say that objects (Jones, the door) enter into the state of affairs which is the content of a belief, it makes no sense to say that objects enter into propositions; names of objects appear in expressions of propositions, but they are not essential to the nature of what is described. Thus consider again "Calcutta is east of Delhi" and "Delhi is west of Calcutta": the relation of being to the east of is named in the former description but not in the latter, though both are expressions of the same proposition. So either we must admit that that relation does not enter into the proposition, or else succumb to the absurd view that all the objects named by any expression of a proposition enter into that proposition, in which case Mrs. Gandhi enters into the proposition presently under discussion, since the expression "Calcutta is east of the city Mrs. Gandhi lives in presently" is also an expression of this proposition. Propositions as classically understood are identical if their truth conditions are identical; the contents of believings-e, by contrast, are only identical if the believings-e are identical. As Wittgenstein noted,[10] Russell appears to have been confused about all this: he writes in Principia Mathematica[11] as if objects did indeed enter into propositions, and as if a proposition were something which might or might not be asserted--thus the introduction of the assertion-sign into his symbolism there. I view his confusion as a symptom of the naturalness of the speech-act model, a symptom showing through in the midst of Russell's construction of an apparatus which is much less natural.

4.4. Meaningfulness in Nyāya is a property of the awareness; in speech-act theory it is a property of the speech-act. For classical analysis it is the proposition which is the primary bearer of meaning; expressions are meaningful if they express a proposition, and a belief is meaningful provided it is a belief in a proposition. But even though the four conditions of meaningfulness specified by Nyāya

appear to correspond with the kinds of conditions classical analysis specifies must be satisfied by a proposition, there is an important difference to be noticed. In speech-act theory these kinds of conditions--expectancy, fitness, contiguity, speaker's intention--are pragmatic conditions governing the successful performance of the speech-act. If I try to perform a speech-act by uttering some sounds and they fail to exemplify these conditions, my attempt is likely to fail to achieve its desired effect. In classical analysis conditions of this sort represent semantic and syntactic conditions, not on the successful performance of an act, but on the meaningful expression of anything, successfully or unsuccessfully. If, to use one of Russell's examples, I say "quadruplicity drinks procrastination" I will not have expressed a proposition at all, so they say. Or again, if I say "and not slips high", I have not expressed a proposition, or if I say "the sky" in 1981, "is" in 1991, and "blue" in 2001, I have not expressed any proposition. Apparently, on the view of the classical analysts, the reason why I have not expressed a proposition is not because I didn't try, but because there is no proposition which has the appropriate structure corresponding to my utterance. The picture is this: there are an indefinitely large number of propositions available which might be expressed, but for an utterance to express them the utterance must satisfy the semantic and syntactic conditions that govern the structure of any proposition. The difference, then, is this: whereas on the classical analysis assumptions an utterance is nonsense just if it fails to express a proposition, on the speech-act theory an utterance is nonsense if it fails to be the sort of act which could function successfully in the way it is intended to. For example, if (as is almost always the case) my speech-act was one of communication (perhaps among other things) then my utterance is nonsense if it fails to be the sort of act which could communicate, e.g., if the form in which it is expressed does not fit the conventions expected by the hearer(s).

4.5. For classical analysis truth is a property of propositions. Beliefs and statements are true just if they have true propositions as their contents. They are false just if they have false propositions as their contents. This leaves open the question what it is about a true proposition that makes it true, or a false one that makes it false. But it does imply that any belief in a proposition is true, or false, solely depending on whether the proposition is true or false, and the truth or falsity of the proposition is independent, antecedent to that of the belief.

What does determine the truth or falsity of a proposition? Not just that it exists, since both true and false propositions exist. Rather, the classical answer has regularly been something like this: true propositions correspond

to actual states of affairs, false propositions to possible but nonactual states of affairs. The contrast between this and the Nyāya position is evident: Nyāya holds that there are no mediating entities between awarenesses and the world, while classical analysis views the proposition as such a mediating entity.

One can spell out the point in a different way without making explicit reference to Nyāya methodological assumptions. The history of the arisal of skepticism in epistemology in modern Western thought shows that the opening wedge leading to idealism and eventual skepticism is when the realist allows that something mediates between awareness and the world. For Locke it was "ideas". For classical philosophical analysis, it is propositions. The nature of the problem is well-known to every student of philosophy: once one allows the postulation of a new mediating entity to answer the question "how is false awareness possible?" one is put in a position such that one cannot defend the existence of anything beyond that mediating entity. It is the mediating entity we are aware of in false awareness; that is evidenced by our being aware of it, and of its not corresponding to anything in the (alleged) real world. But if we can be aware of things which are not in or do not correspond to the (alleged) actual world, we have no way of proving that what we are aware of in true awareness is not likewise a mediating entity. Thus idealism and/or skepticism hold sway. Any system which hopes to maintain epistemological realism (and not just Nyāya, which is only one such system) should therefore look askance at introducing propositions in the way that classical analysis does.

4.6. In speech-act theory, we saw, the content is identified through the act, and in parallel fashion, in Nyāya the content of a mental act is identified through the act. In classical analysis it is the other way around; an expression uttered on a given occasion expresses the proposition that is expressed by that expression, regardless of the intention of the speaker. This affects both the neustical and phrastical aspects of an expression: classical analysis treats declarative sentences as expressive of propositions even when they are not being used to assert. E.g., even though by my saying "the door is open" I intended my speech-act to have the force of a request or command (tantamount to "shut the door"), the classical analytic philosopher will still treat my remark "literally", as the statement of the proposition _the door is open_. Again, if I assert _p_ by using an expression standardly understood as meaning _q_, the proposition I have expressed is _q_, not _p_.

4.7. As a result, it is impossible for someone, given classical analytic assumptions, to say one thing and mean another. Furthermore, though there is a sense (or perhaps

more than one sense) that can be attached to the notion of the same proposition falling under different descriptions, it is not the same sense as the one used in Nyāya and natural to speech-act theory. Whereas we can say, consistently with the speech-act assumptions, that S said "q" intending p as his phrastic, even though "q" is conventionally understood by hearers as meaning q, and thus S might be described as referring to p under the description "q", this cannot be the case of traditional analysis. What I refer to is not a matter of my intentions for traditional analysis: it is a matter of what words I do in fact use. I might try to mean p by saying "q", which is to say I might intend to be understood that way, but what I referred to by uttering "q" was q.

5.

In recent years a number of us have been writing in English on Navya-Nyāya and Indian logic generally, and the term "proposition" has been bandied about freely. Sometimes these treatments show ambiguity and/or confusion about the notion(s), as in S.S. Barlingay's A Modern Introduction to Indian Logic, where a lot is said about "the Nyāya theory of propositions", but it never is unambiguously clear what is being intended by that expression. At some points it seems to be used synonymously with "assertive judgment",[12] at others we are told that a proposition is different from a judgment,[13] at still others the proposition is the predicate of the judgment,[14] and at still others propositions are treated in the classical analyst's manner as bearers of truth-values.[15]

In 1966 J.N. Mohanty published his classic Gaṅgeśa's Theory of Truth, in which he distinguished carefully between judgments (jñāna) and propositions. Nevertheless, he proposes to term savikalpaka jñāna "propositional", arguing that "it is a logical complex analysable into constitutent elements and relations."[16] In Mohanty's review[17] of B.K. Matilal's The Navya-Nyāya Doctrine of Negation[18] he pursues an analogy between phenomenology in Brentano and Husserl and Navya-Nyāya's way of treating acts of awareness and their contents. The analogy proves helpful but eventually breaks down because Nyāya insists on the reality of both awareness and content, whereas the phenomenological doctrine of intentionality is "ontologically neutral". Nevertheless, it may be that the phenomenological model is as good or better than the speech-act model discussed here: my point is not that the speech-act model is the only or even the best model, but only that it is a superior model to classical analysis in terms of propositions.

Matilal's book is admirably attentive to the distinctions between propositions and acts of awareness, as well as to many of the points made in the foregoing sections of this paper.

Nevertheless, he is taken to task by J.F. Staal in his review[19] of Matilal's book for employing a symbolism which obscures the difference between a proposition and a term. What Staal takes to be a "proposition" is "x is qualified by y", in contrast to the term "x which is qualified by y", and the representation of the content of a jñāna then turns out to be a proposition for Staal. In responding to this, Matilal alludes to Frege's distinction which we have termed that between the neustic and the phrastic. Matilal writes, "What I have done in my (1968) [i.e., the Negation book] is to take some liberty with Frege's method of analysis and propose that if we take the assertion sign out, the content of the proposition is reduced to a COMPLEX of terms representing a state of affairs."[20] This is part of the move I am recommending, away from the propositions of classical analysis and toward the phrastics of speech-act analysis.

The most recent development in this continuing critique of analytic analogies for Navya-Nyāya comes in Sibajiban Bhattacharya's "Some Principles and Concepts of Navya-Nyāya Logic and Ontology."[21] Bhattacharya also chooses Frege for his point of departure, but concentrates on the point that in Frege the sense of a proposition represents the mode of relation in which the constituent terms stand to each other, whereas in Navya-Nyāya that mode of presentation of the terms can only be shown, not represented. "A sentence is incapable of expressing the manner of presentation of the entities denoted or designated by the terms; the moment the manner of presentation is sought to be expressed in words, they themselves become nominata and their manner of presentation has again to be understood. The defect of Frege's theory, according to the Navya-Nyāya, would be that although Frege wanted to explain the 'different cognitive significance' of sentences, he still did not pay any attention to the cognitions, and devoted himself exclusively to an analysis of sentences."[22]

6.

I have reviewed the literature to suggest that other writers, including some of the most careful and diligent scholars of Navya-Nyāya who are also thoroughly aware of developments in analytic philosophy, are growing uneasy with the model of classical analysis, either in Russell's or Frege's version, for appreciating the precise thrust of Navya-Nyāya. No doubt they, capable of appreciating the full nuances of Sanskrit usage, do not feel any great need for supplying a replacement to these analogical models. However, where the problem is one of expounding Navya-Nyāya to a Western audience of students and scholars of philosophy unacquainted with Sanskrit,[23] the question of an expository model becomes relevant. My

suggestion is that by extending speech-act theory in a natural manner one can arrive at a model quite suitable for this purpose.

Let me conclude by setting down succinctly some of the most obvious features of this extension together with a terminology in which to express it.

6.1. Austin[24] identifies the speech-act as the fundamental element of meaningful language. Extending this to epistemology, I suggest that the awareness act (jñāna) is the fundamental element of cognition.

6.2. Austin and others[25] have identified and classified a large number of varieties of speech-acts, which can also be termed varieties of neustical aspects of speech-acts. Austin classifies such aspects into five large groups, viz., verdictives, exercitives, commissives, behabitives and expositives. These have also been referred to as kinds of illocutionary force. Matilal lists a number of types of jñāna thus: "Perceiving, inferring, knowing, doubting, wondering, guessing, remembering, dreaming, etc."[26] These, then, are samples of the neustical aspects of awareness-acts.

6.3. Austin[27] distinguished the "locutionary force" of a speech-act from its "illocutionary force," apparently intending to allude to what we have been, following Hare, calling the "phrastic". Austin doesn't have a lot to say about the kinds of questions we have been concerned with here having to do with the constitution of the phrastic; strategically, for speech-act theorists the point of speech-act theory is better made by concentrating on the act rather than its content. Still, speech-acts clearly do have contents (although not all of them may have contents with structures, as Searle[28] notes), and so do awareness-acts have contents (viṣaya), although not all awareness-acts have contents with structure either, e.g., nirvikalpaka jñānas in Nyāya.

6.4. Speech-acts have speakers and hearers (usually, perhaps even always if the speaker can also be the hearer of his own speech-acts). Likewise, awareness-acts have agents (jñātṛ).

6.5. Truth and falsity are attributes of speech-acts, not of their phrastics, in speech-act analysis of Austin's sort. (He argued with Strawson over this point.[29]) Likewise, Nyāya treats truth and falsity as attributes of awareness-acts, not of their contents.

6.6. Austin distinguishes locutionary and illocutionary aspects of speech-acts from a third aspect, that of "perlocutionary force", which has to do with the intended result(s) of a speech-act, what one is intending to accomplish by speaking. Likewise, we can distinguish the purpose (prayojana) of an awareness-act. Whether all acts--of speech or of awareness--have this kind of force is a matter for discussion; it is already being discussed in the literature on speech-acts.

6.7. The product of a speech-act is distinct from the process of speaking, i.e., producing it. Strawson locates meaningfulness in the products of speech-acts, in statements.[30] (This puts him at odds with Austin, and also more in sympathy with classical analysis than our own predilections as offered here dictate.) One can also distinguish commands from commanding, warnings from warning, and so on. Likewise, one can distinguish the product of an awareness-act--e.g., a thought, an idea, a belief--from the having of the thought, etc., from the awareness in the sense of the act. (Sanskrit distinguishes the pramiti from the pramāṇa.)

6.8. Finally, a somewhat controversial point in speech-act analysis, which relates to the importance of H.P. Grice's work[31] in the area. In identifying which speech-act (in either its neustical or its phrastical aspect) is being performed one can take one's cue from the expressions used in the utterance or from the speaker's intention. So, on the former reading the meaning of a speech-act stems from the conventional meanings of the constituent elements (with problems arising concerning nonliteral uses, which certainly occur), whereas on the latter reading the meaning stems from the intention of the speaker (so that nonliteral uses pose no problem, but a different problem arises, how to tell what a speaker means). The distinction is close to Grice's between natural and nonnatural meaning.[32] Speech-act theory has tended to develop to some extent along the lines of the latter reading, though with reluctance and opposition expressed in certain quarters. Likewise, one could raise the question for awareness-acts whether they are to be identified through their agents' intentions or in some other fashion. It is hard to make a very plausible case for any other fashion here, however, and it is no surprise that Nyāya invokes speaker's intention (tātparya) as an individuating feature of jñānas, though it is interesting to note that allusion to intention is a controversial matter in Nyāya discussions also.

This point is connected with another. Are all awareness-acts in effect expressed by language, so that talk about awareness-acts is necessarily talk about speech-acts, at least about what we would say if we chose to articulate our thoughts? Or is thought an activity independent of linguistic activity? The Naiyāyikas say they are distinct.[33] The interesting point is that the discussion arises in these terms there, which gives additional support to the thesis maintained here, which is that the awareness-act theory is not identical with, but an extension of, the speech-act theory which provides a happy model for English language expositions of Navya-Nyāya, one preferable to classical analytic philosophers' ways of treating these issues.

University of Washington

NOTES

1. E.g., Bimal Krishna Matilal: 1968, The Navya-Nyāya Doctrine of Negation (Harvard University Oriental Series 46), Harvard University Press, Cambridge, Mass, p. 46; J.L. Shaw: 1980, 'The Nyāya on Cognition and Negation', Journal of Indian Philosophy 8.3, 289-299; Karl H. Potter: 1970, 'Realism, Speech-Acts, and Truth-Gaps in Indian and Western Philosophy', Journal of Indian Philosophy 1.1, 13-21.

2. E.g., Kali Krishna Banerjee: 1958, 'Wittgenstein versus Naiyāyika', Calcutta Review 147, 27-44; Kisor Kumar Chakrabarti: 1975-76, 'Some Comparisons Between Frege's Logic and Navya-Nyāya Logic', Philosophy and Phenomenological Research 36, 554-563; Kalyan Kumar Sengupta: 1969, Language and Philosophy, Allied Publishers, Bombay; Kaisa Puhakka: 1975, Knowledge and Reality: A Comparative Study of Quine and Some Buddhist Logicians, Motilal Banarsidass, Delhi; Hemanta Kumar Ganguli: 1965, Philosophy of Logical Construction, Calcutta.

3. R.M. Hare: 1952, The Language of Morals, Clarendon Press, Oxford, pp. 17-21.

4. R.M. Hare: 1948, 'Imperative Sentences', Mind 58, 21-39.

5. John R. Searle: 1969, Speech Acts, Cambridge University Press, Cambridge, pp. 29 ff.

6. Sometimes this fourth is omitted. The significance of this is discussed later in this paper.

7. This theory of error is known as anyathākhyātivāda. The point is that the terms which appear in false judgments as well as true ones all designate actual entities; the falsity of the judgment is not a function of the status of what its terms designate, but of what the judgment shows (rather than says) about them. See below for further comment on this feature of Nyāya.

8. Hare, Language of Morals, op. cit.

9. E.g., Searle, op. cit. as well as Stephen R. Schiffer: 1972, Meaning, Clarendon Press, Oxford, tend to go to the other, more classical, way on this point.

10. Ludwig Wittgenstein, Tractus Logico-Philosophicus 4.442; also 3.143 and 5.02.

11. Alfred North Whitehead and Bertrand Russell: 1962, *Principia Mathematica to *56*, Cambridge University Press, Cambridge, pp. xix- 8, 43, etc.

12. S.S. Barlingay: 1976, *A Modern Introduction to Indian Logic*, 2nd revised and enlarged edition, National Publishing House, New Delhi, p. 46.

13. Ibid., p. 45.

14. Ibid., p. 51.

15. Ibid., Chapter 5 ff.

16. Jitendranath Mohanty: 1966, *Gaṅgeśa's Theory of Truth*, Centre of Advanced Study in Philosophy, Santiniketan, p. 31.

17. In *Journal of Indian Philosophy* 1.2, 1971, 197-211.

18. Matilal, op. cit.

19. In *Indo-Iranian Journal* 13.3, 1971, 199-205.

20. Bimal K. Matilal: 1971, *Epistemology, Logic, and Grammar in Indian Philosophical Analysis*, Mouton, The Hague, p. 92.

21. Sibajiban Bhattacharya: 1976, 1977, 'Some Principles and Concepts of Navya-Nyāya Logic and Ontology', *Our Heritage* 24.1, 1-16, and 25.1, 17-56.

22. Ibid., pp. 25-26.

23. My own interest in this question is made especially urgent by the problems presented to myself and my co-editor Sibajiban Bhattacharya as we try to develop an appropriate expository vehicle for our forthcoming volume in *The Encyclopedia of Indian Philosophies* which will deal with Navya-Nyāya from Gaṅgeśa to Raghunātha.

24. J.L. Austin: 1962, *How to Do Things with Words*, ed. J.O. Urmson, Oxford University Press, New York.

25. Austin, ibid., pp. 147 ff. See also papers by Zeno Vendler and others proposing improvements on Austin's lists, e.g., in Vendler's 1972, *Res Cogitans*, Cornell University Press, Ithaca, N.Y.

26. Matilal, *The Navya-Nyāya Doctrine of Negation*, p. 8.

27. Austin, op. cit., p. 99.

28. Searle, op. cit., p. 30.

29. Cf. 'Truth', Proceedings of the Aristotelian Society, Supplementary Volume XXIV, 1950, also in George Pitcher, ed.: 1964, Truth, Prentice-Hall, Englewood Cliffs, N.J., pp. 18-53.

30. Strawson in 'Truth', ibid.

31. H.P. Grice: 1957, 'Meaning', Philosophical Review 66, 377-388.

32. Grice, ibid.

33. See Sibajiban Bhattacharya, op. cit., pp. 29-31.

Douglas D. Daye

SOME EPISTEMOLOGICALLY MISLEADING EXPRESSIONS: "INFERENCE" AND "ANUMĀNA", "PERCEPTION" AND "PRATYAKṢA"

1. INTRODUCTION

In this article, I shall focus on the four terms of the title. The implications of the present remarks will be argued elsewhere.

By the term "AE logics", I refer to a cluster of formal machineries common to the Anglo-European tradition of formal logic which range from syllogistic in its varied forms and uses to the first order predicate calculus as used by Nyāya scholars of this century. Coupled with these, are inter-webbed philosophies of logic and a whole variety of epistemological, ontological and metaphysical doctrines and assumptions. By the terms, "PA, Buddhist logic," etc., I refer to the Indian Buddhist pramāna vāda doctrines regarding "anumāna, parārthānumāna (cited hereafter as the "PA") svārthānumāna" and an analogous (and only analogous) not isomorphic, cluster of epistemological, ontological and metaphysical doctrines and assumptions. In particular, I focus here on the (so-called) "inference-schema" of the PA and some allegedly relevant and/or compatible AE notions about formal logic(s); I note the very wide-spread use of the latter as a formalistic target ideal language for the translation of the PA via the AE logical machineries.

The AE and PA formal/formalistic logics are only the tips of one iceberg of cognitive knowledge; formal logic is one type of reliable human knowledge. Our attitudes about and approaches to knowledge, are exceedingly varied and complex. Also the "objects" of our inquiries are so complex, so intertwined, so varied, that the possible constellations of globally oriented approaches and problems inherent in examining Buddhist "perception," remain enormous. However, before I review the relevant texts, let me state my views of the problems to be examined.

This article remains heuristic rather than conclusive, for the focus is logic, not epistemology.

What I wish to show here is that the range of things which these early Indian Buddhist logicians meant by the term "anumāna," is not fully commensurate with what is meant when most (if not all) twentieth century AE oriented Nyāya

B. K. Matilal and J. L. Shaw (eds.), Analytical Philosophy in Comparative Perspective, 231–252.

scholars use the English target language term "inference." Further, I will claim that translating "anumāna" as "inference," misleads the general reader and skews their expectations of the PA. Thus, in this article I assert a negative thesis. I do not try to solve enormous problems, which would take me far beyond the scope of this paper, as to what either the full range of Indian or Indian Buddhist answers were, or what "the" or "a" legitimate set of answers to these thorny problems of translation and analysis of epistemology and ontology should be. Rather, I wish to show the incorrectness of the case for using artificial languages (= AL): 1) as the media of translation for the PA; 2) as the fundamental model of interpretation to which, it has been held, that the Buddhist logicians were aspiring; and 3) as the model to which they should have aspired. At least, such uses constitute a questionable philosophical projection on "anumāna" as in PA, and thus may be misleading for the scholar of comparative formal logics. Hence, I shall consider "anumāna" insofar as it bears on these problems of comparative logics. I now turn to some aspects of this problem(s).

First, a generalization is in order; contrary to the usual, the AE sense of (logical) propositions, neither the PA sentences nor propositions were held by the Indian logicians to be bearers of truth-values. In general, "cognitions" (jñāna) were the bearers of what AE logicians would call truth values.

Second, Buddhist philsophers in general conceived of anumāna, both for one's self (svārthānumāna = SV) and anumānas for others (the PA), by means of two general models of causality (kārana) and/or of identity (svarūpa) relationships, but not by means of the intertwined AE models of the analytic argument and the deductive concept of formal validity.[1] This is not to be taken as a claim that no proto-analytic arguments can be found in the PA texts, for as is demonstrable, (almost) all of the PA models in the Nyāya Praveśa (NP) are necessarily true once the suitable transformations of reversing the order between the thesis (pakṣa) and the justification (hetu) and warrant (dṛṣṭānta) have been implemented. However, as I demonstrate elsewhere, neither the use of nor the concept of any AE formal (deductive) validity is present in the justification procedure of a specific PA. The general formalistic tendency in Nyāya texts is present, but the process of comparing the formal pattern of the PA to abstract models of deductive rules of valid inference pattern is not present.

Karl H. Potter succinctly notes this point regarding the Nyāya darśana:

> Thus where we conceive of the validity of an inference as compatible with the failure of its members to refer or be true of anything, the Naiyāyika views nonreferential words as ill-formed and excludes them from any inference.
>
> This fact seems to me to show that Nyāya is not concerned with "formal" logic in the way that Western logicians have characteristically been. However, this does not contradict Bochenski's view that "in India too a formal logic developed." That it really was a formal logic is shown by the fact that the formulae constructed by the Indian thinkers concern the fundamental questions of logic, the question of "what follows from what..."[2] Of course Indian logicians were interested in what follows from what. But their logic was not "formal" in a different sense; it did not have to do with abstract relations among terms, where the abstraction was[3] from all questions of reference. A Western logician views the inference "all animals are pigs; all pigs have wings; therfore all animals have wings" as formally valid, though unsound. The unsoundness does not in his view, detract from the logical interest of the example as instantiating a valid form of inference. The Naiyāyika's view of this example is that it is a nyāyābhāsa, something which is only apparently an argument but really is not. It is, in short, ill-formed because its members are known to be false.

Thus for such an Indian PA to be illegitimate is not grounded in the familiar AE distinction between the truth of the premises and the formal validity of the argument; clearly the PA fails to exhibit such a concept of formal validity.

Third, most AE oriented scholars of Indian logic, on seeing "anumāna" as "inference", become unreflective absolutists exhibiting very low tolerance for the further examination of the epistemological complexities of anumāna vs. pratyakṣa with respect to the criteria of an alleged formal logic. It is also true that some AE logicians are reluctant to recognize the complexities and examine the underlying epistemological presuppositions regarding the use of artificial language expressions as target languages for natural language proto-arguments; however, AE logicians are generally more sympathetic to the need to examine such questions in philosophical logic than are most Buddhologists. These latter epistemological problems comprise the huge mass (of concepts) shouldering the

tip of the iceberg of comparative formal logics. For three illustrations, note the emic AE controversies regarding the "paradoxes" of material implication; for an etic example (with Nyāya), note the widespread use since the 1950s of truth-functional AE operators ("⊃") with which to translate and describe the PA ablative case in the hetu; whether the conditional of the dṛṣṭānta is compatible with material implication is a third controversial feature of comparative philosophy of logic.

Hence it will be necessary to ask questions about: 1) whether the natural language PA expressions are actually bearers of truth-values; 2) whether the reversal of the ablative case into an AE "standard form" with the "premises" (hetu, dṛṣṭānta, sapakṣa and vipakṣa) first and the "conclusion" (pakṣa) last is misleading, or 3) whether the AE analytic argument and the deductive concept of formal validity are what the Buddhist pramāṇa-vādins really had in mind when they discussed the legitimacy of a PA. However, the two logically prior questions are: what is anumāna, and what kind of anumāna is PA?

The unrecognized problem of the allowable degrees of complexity regarding the cross-cultural translations of the PA emic metalogical concepts are significant whether or not its wholesale translation into AL expressions is misleading. However, a generally unacknowledged problem remains within Indology regarding both the PA and the epistemological assumptions about "anumāna" as "inference" which underlie its emic development and analysis. What is usually missing is an awareness of the recognition of the epistemological and (thus) metalogical complexity of such Nyāya translations of both the PA as "inference-for-the-sake-of-others" and "anumāna" as "inference."

Fourth, if "inference" is not an altogether fitting translation of "anumāna," what English word(s) might be used? A definitive answer to this question would require both a different and a more inclusive epistemological study than I am now able to offer. Thus, I do not now offer a final analysis; I merely wish to offer a few cautious, heuristic suggestions regarding the general problem, to convey controversial features of its context and to offer some suggestions for new translations.

2. THE TEXTUAL SOURCES

To ground my discussion, let us now turn to the nyāya texts. The word "anumāna" is always related, in the Buddhists pramāṇa-vāda texts, to the term "perception" (pratyakṣa). The term "anumāna" has been (almost) always translated as "inference". It is one of the contributions of the early Buddhist logicians (probably Dignāga), that they/he divided the

pramāṇa of anumāna into two kinds, namely "inference for one's self" (svārthānumāna = SV) and "inference for others" (parārthānumāna, the PA). At first glance, we recognize the latter kind of "inference", the PA, as akin to a fairly common inference making pattern in deductive AE logic, Modus Ponens.

I defer an examination of the latter point, but I would first claim that the word "inference" does not fit terribly well for either what is meant by "inference for one's self", the SV, or for the pramāṇa of anumāna. Thus let us examine the definitions and relations between anumāna and pratyakṣa.

The NP offers a very brief statement. "Perception (pratyakṣam) is devoid of conceptual construction (kalpanāpoḍham). It is that type of cognition (jñāna) which does not superimpose or construct any name, generic category upon an object of perception such as color, etc. (Perception) is called by such a term because it occurs to (prati) each object and the power of the sense organ (akṣa). Inference is the understanding of an object through its identifying characteristic. We have noted that the identifying characteristic has three aspects."[4]

The Chinese translation of the Nyāya-mukha (NM) states the following. (This is a modified translation following Tucci and the Taishō.

> With regard to the following sentence, "the other (anumāna) is derived from the justification already explained," it is necessary to understand the (following) concept of "knowledge." This (pratyaksa, "perception") is the first type of knowledge. The other type (anumāna, "inference") is derived from the justification-justifier relations which are part of the demonstration, as explained, and is based upon the significance of the relations (of that demonstration). (This knowledge) can be of two kinds, that is, the knowledge which consists in the apprehension of the object to be known (sādhya) and is derived from either perception (pratyakṣa) or from "inference" (anumāna) plus the recollection of the invariable concomitance between the justifier (hetu) and the thesis (pakṣa and the property-to-be-demonstrated, sādhya). Hence there is demonstrated the thrust of the PA (assumed before) because there is the memory (of those related characteristics of the trirūpahetu which gives legitimacy to the justifier), that the justification actually occurs in the similar exemplification, (sapakṣa) etc.... it is necessary hence to be

> cognizant that the PA also cannot remain a legitimate means of demonstration (knowledge) without also assuming the legitimacy of the process of making an "inference for one's own self" (svārthānumāna).[5]

Thus to paraphrase, it is clear that one cannot attempt to demonstrate the legitimacy of those sui generis epistemological characteristics, of momentary instantaneous raw perception (pratyaksa), without also presupposing the use of both the machinery of the PA, the SV and particularly the "three forms of the justifier-property" (trirūpahetu).

3. SOME SECONDARY STUDIES

We now turn to an exposition from Nyāya scholars of the definitions, characterizations and the constellations of relations regarding the terms "pratyakṣa", "anumāna" and the dual sub-set of the latter pramāṇa, svārthānumāna (= SV) and the PA. As the texts of the NP and the NM have relatively little discussion of pratyakṣa, I turn to three classic secondary sources, Sanghavi, Matilal and Hattori.

The late Pundit Sanghavi discusses "pratyakṣa" as follows:

> In Buddhist Logic there are two traditions of defining 'perception'--one which does not include the word "non-illusory" (abhrānta) in the definition, the other which does. The first tradition is initiated by Diṅnāga, the second by Dharmakīrti. Thus the definition and description given in Pramāṇa-samuccaya (1.3) and Nyāyapraveśa ... follow the first tradition, those given in Nyāyabindu (1.4) and its commentary by Dharmottara etc. [follow the second].[6]

B.K. Matilal characterizes pratyakṣa in the tradition of Dignāga, the author of the NM and a rough contemporary of the relatively unknown author of the NP, Śāṅkarasvāmin.[7]

> Diṅnāga defines perception as a cognitive state which is totally untouched by imaginative construction (vikalpa, kalpanā) or conceptualization. He accpets the view implied by Bhartṛhari that conceptualization and verbalization, construction and language, are just two aspect of the same process. Construction, Diṅnāga says, is nothing but our associating any name, viz., a proper name or a class name or a

quality-name or an action-name or a substance-name, with the datum. In other words, it is an imposition of our constructive forms upon the 'given.'

In each case, our judgment (let us call it judgment instead of construction in this context) implies a positive relation or a relation of qualification. Thus, from the popular point of view we can call the five types of names mentioned above the five designators, and very loosely, the five predicate-constructions. Diṅnāga calls them the five 'qualifiers' (viśeṣaṇa). In fact, the function of each name, each designator, according to Diṅnāga, is to qualify, and to qualify means to differentiate. Qualification and distinguishment or differentiation are, for Diṅnāga, just different names for the same function.

Diṅnāga illustrates the five 'qualifiers' as follows: In the case of the proper names such as Diṭṭha the thing appears as qualified, or distinguished, by the name itself. In the case of a class name such as cow the thing appears as qualified by a class concept, cowness. In the case of a quality name such as white it appears as qualified by a quality concept, white color. In the case of an action name such as cooks or a cook (pācaka) the thing appears as qualified by an action concept, cooking. In the case of a 'substance-name' such as a staff-bearer (dandin) or horned (animal) (viṣāṇin) the thing appears as qualified by a substance concept, a staff or horns.

Diṅnāga is careful to point out that according to some the 'qualifiers' are not such 'concrete' notions as a staff or horns and the action of cooking, but rather the abstract properties like the thing's relation to the staff or horns and the agent's relation to the action .... Diṅnāga emphasizes the 'quality' aspect of meaning and hence focuses attention upon the qualifiers or 'distinguishers' and neglects completely what is purported to be qualified or distinguished. Unlike the Universalists (e.g., Vājapyāyana) he does not want to hypostatize the qualifiers ... all these qualifiers distinguish their objects by virtue of being MERE names of them without implying in any sense real entities.

> All the qualifiers in Diṅnāga's list belong to what Diṅnāga calls the category of sāmānya, universal or generality. To call something a universal is, in Diṅnāga's system, to say that it is an imaginative creation, a conceptual construct, and it must be sharply distinguished from the given or the datum, which Diṅnāga calls the 'exclusive particular' (sva-lakṣaṇa). The given, in Diṅnāga's system, is always characterized by a uniqueness of its own, which is not repeatable and also not definable or expressible in language. Only the unutterable, unrepeatable, and unique particulars are reals, the universals are at best unreal superstructures and at worst fictional images.
>
> Perception, i.e., perception proper, according to Diṅnāga, is revelation of the pure given, the unique datum. It is untouched by construction and hence uncontaminated by any sort of speech or language whether implicit or explicit.[8]

Matilal summarizes "perception" (pratyakṣa):

> Thus perception, being mere revelation of the given and entirely free from our subjective manipulation, is completely reliable and has an absolute truth-claim in Diṅnāga's system. The truth claim of other cognitive states is relative. The possibility of what we may call perceptual illusion does not apparently clash with this absolute truth claim of perception as far as Diṅnāga is concerned. Whenever the mind is at work conceptualizing there is room for error. But if no mental processes are at work, there can, ipso facto, be no possibility of error. Thus according to Diṅnāga's theory, to use the expression 'perceptual error' is to misue the term 'perceptual' which is defined in such a way as to preclude error.[9]

M. Hattori supplies further support by stating that "according to Dignāga, a thing, which in itself is essentially inexpressible, comes to be expressed by a word only when it is associated with a name (nāman) and other factors. Conceptual construction (kalpanā) means nothing other than this process of associating a name, etc., with a thing. Dignāga classifies the factors to be associated with a thing for the sake of verbal designation into five categories: nāman,

jāti, guṇa, kriyā, and dravya."[10]

Thus "pratyakṣa" (perception) does not name any "real objects" in any usual sense of "object," for this would be to beg the ontological questions at issue between the Buddhists and their opponents. Rather, "pratyakṣa" is a name for purported "things" which are non-anumāna. The necessary conceptual distinctions to make this provisional division are superimpositions projected upon a fluxing, ever changing, instantaneous multitude of svalakṣaṇa(s), "things" (in the loosest sense); these are allegedly unique, flashing staccato bits of experience. These can never be adequately described by generic words or names, for all presuppose to some degree, the non-uniqueness, the commonalities presupposed in the use of universal terms; these are words which refer to an alleged something which possesses qualities perceived as sufficiently similar, to be grouped together and identified for purposes of alleged reference and intersubjective communication. This communication about "reality" is possible with and about anumāna, but not about the flux we call "pratyakṣa." We know that anumāna cannot be pratyakṣa, but, is it "inference"?

On "anumāna," Sanghavi offers a general summary description.

> The word "anumāna" (i.e. inference) means two things, viz. inferential cognition (anumiti) and the instrument of inferential cognition (anumitikaraṇa). Thus when the word stands for an abstraction (bhāvavācin) it means inferential cognition, when it stands for an instrument (karaṇavācin) it means instrument of inferential cognition.
>
> The word "anumāna" consists of two parts, viz. "anu" and "māna". "Anu" means 'after' and "māna" means 'cognition,' so that "anumāna" means 'a cognition taking place after some other cognition'. But this other cognition has to be a particular type of cognition, a type which alone acts as the cause of inferential cognition, and 'cognition of pervasion' (vyāptijñāna, cognition of invariable concomitance), otherwise known as 'consideration of the probans' (liṅgaparāmarśa, *"consideration of the justifier-property") is the type in question. One outstanding difference between perceptual cognition and inferential cognition is that the former is not necessarily caused by another cognition while the latter is necessarily so caused; it is this idea that is conveyed by the part "anu" present in the word "anumāna". Although there are certain types of

> non-perceptual cognition--e.g., cognition through analogy (upamiti), cognition through verbal testimony (śabda), and cognition through implication (arthāpatti)--which are generally not treated as the cases of inferential cognition, the fact of the matter is--and the Vaiśeṣika and Buddhist systems recognize it--that the pramānas are of only two types, perceptual and inferential. As for the remaining types of non-perceptual pramāṇa, they can all be somehow treated as cases of inferential pramāna--as has been done by the two systems just referred to.
>
> Whatever be the object of a piece of inferential cognition and whatever the type of probans (*justifier) causing it, it is definite that all such cognition must have a piece of perceptual cognition somewhere at its basis. For an inferential cognition having no perceptual congition somewhere at its basis is an impossibility. Thus while perceptual cognition comes into existence without at all depending on inferential cognition, inferential cognition comes into existence only in dependence on perceptual cognition. It is this ideal that has been expressed by the Sage (ṛṣi) Gotama through the phrase "tatpūrvakam" (i.e. preceded by it, that is, by perceptual cognition) occurring in the defintion of inferential cogniton given by him in Nyāya-Sutra (1.1.5.).[11]

In Hattori's study of Dignāga's theory of pratyakṣa, he states:

> Anumāna... literally means a means of cogntion which is preceded by some other cognition. According to the Naiyāyikas, that which precedes anumāna is perception of a mark (liṅga, *justifier-property, hetu) and of the invariable connection between this mark and its possessor (liṅgin, *property-possessor, dharmin) .... Thus, the prefix "anu-" is taken by the Naiyāyikas to mean "paścāt" (afterwards) or "-pūrvaka" (preceded by); ... Dignāga, however, interprets differently the meaning of "anu-". His definition of anumāna for one's own self (svārthānumāna) is: ... (That apprehension of an object which is based upon the triple-conditioned inferential mark is svārthānumāna); ... Since Dignāga regards determinate

perception (savikalpaka pratyakṣa), which perceives a thing as associated with a universal (jāti-viśiṣṭa-vyakti), as a kind of anumāna, the terms "pratyakṣa" and "anumāna" in this treatise are to be understood as standing respectively for 1) direct, unmediated cognition or immediate awareness and 2) indirect, mediated cognition. In translating, for the sake of convenience, I employ the term "perception" as an equivalent for pratyakṣa, and "inference" for anumāna.[12]

4. ANALYSIS OF "PERCEPTION" AND "PRATYAKṢA", "INFERENCE" AND "ANUMĀNA"

Pratyakṣa is, then, a direct unmediated cognition of an ever-changing flux, which may subsequently become discursively indirectly known as what we usually call in ordinary language (cited hereafter as OL), a perception. That is, what we usually mean in OL is a "perception of X and not of non-X," thus presupposing the very discursive distinctions outlawed in the definition of pratyakṣa. On the other hand, anumāna is that type of cognition which possesses those very discursive distinctions presupposed in our OL term "perception". Let us compare the two. Pratyakṣa is raw, conceptually unfiltered "experience," flashing flux (svalakṣaṇa); anumāna is mitigated, structured cognition in which the binary relationship of cognized concomitance (vyāpti-jñāna) is the basis of the inter-webbed distinctions. These distinctions are rooted in the filter of generic terms and relations (sāmānya-lakṣaṇa). It is the function of the proto-rule of the "three forms of the justifier-property" to establish the evidence for (necessary conditions of) the filter of concomitance (vyāpti).

In short, we, as all too human individuals, cannot ever directly "grasp" the fluxing, staccato pratyakṣa: we cannot do so except insofar as we superimpose on pratyakṣa, generic words and names accompanied by their webs of intra-defined relationships. When we communicate about pratyakṣa, we are using the generic terms of sāmānya-lakṣaṇa; we are using anumāna, not pratyakṣa. Insofar as we know something, as a cogntion (jñāna), we are "in" anumāna, not pratyakṣa. Thus anumāna is the basis for our everyday experience. When we ordinarily "perceive", when we know something as a legitimate cogniton, we recognize it by means of the superimposition of the generic anumāna machinery.

Hence in anumāna, we have a metaphysical-ontological doctrine about the epistemological foundations of Buddhist logic. Logic, in its public PA articulation, examines strings of words, through which we publicly discern cognitions

(jñāna), the epistemological foundations of which contrasts anumāna, the conceptual locus of logic, with an articulated theory of ontology. For Buddhists, this generic-sounding ontology is at best a distortion of "what is really there". Let us turn to the criteria of anumāna.

First, for a cognition of anumāna to be legitimate, there is the required satisfaction of the proto-rule, the trirūpahetu, the "three forms of the justifier-property" (hetu-dharma). These criteria for all anumāna are taken from the development of the PA machinery itself; it is also a necessary condition of SV too. Anumāna must satisfy that: 1) the property of the thesis (pakṣa dharma) must be present in or concomitant with the property of the hetu, "reason" or as I prefer, the "justifier" of the "justification," the second "member" (avayava) of the PA; 2) the allegedly concomitant thesis-justifier properties (pakṣa dharma and hetu dharma) must be present in or concomitant with the properties of the similar exemplification (sapakṣa) attached to the explicit statement of concomitance of the warrant (dṛṣṭānta); and 3) neither the thesis-property (pakṣa dharma) nor the justification-property (the hetu, justifier dharma) can be present in or concomitant with the properties of the dissimilar exemplification (vipakṣa).

Hence the concomitance relations of anumāna are binary; that is, they are either present or nonpresent. Pratyakṣa, to the extent to which it can be said to be non-privately known at all, is known via anumāna. Pratyakṣa is 1) mitigated or differentiated by means of the superimposition of generic terms, names, relations (sāmānya-lakṣaṇa), by conceptual construction (vikalpa, kapanā), and 2) is characterized by superimposition by the five "qualifiers" (viśeṣaṇa, as noted above in Matilal).

However, with both the SV, the "inference-for-one's own sake" (sva), and the PA too, the trirūpahetu rule applies as a necessary condition. It can be demonstrated that in the PA there is no concept of deductive validity, no truth-preserving transformation rules (such as $(p \supset q) - (nq \supset np)$), nor is the actual mode of justification that a PA is acceptable, similar to the (only slightly) analogous AE procedure of justifying a true deductive inference as a sound inference. Common AE features of justifying a true deductive inference and its crucial machinery (formal validity) are not found in the PA. If not in the PA, then, a fortiori, what of all our ordinarily experienced knowledge which is not that of a PA? That is, what of the SV, the so-called "inference-for-one's-own-sake" or "inference-for-the-purpose-of-one's-own-self"?

To examine this last question, certain assumptions need to be made explicit. For these Buddhist "logicians", the legitimate sources and means of knowledge (pramāṇa) are pratyakṣa and anumāna. Ordinary experience and ordinary

language and the language of epistemology and ontology employ generic terms, etc. By definition such language cannot accurately express the flashing particulars (svalakṣaṇa) of pratyakṣa. We are thus left with anumāna for which ordinary philosophical language is (allegedly) *sometimes* adequate. The chief criteria for legitimate anumāna expressions are the satisfaction of the "three forms of the justifier" (trirūpahetu); this requirement holds for both the SV, "for-one's-sake" and the PA, "for-the-sake-of-others".

This latter distinction, I assume, is directly identical with the standard AE distinction between first person claims or language and third person claims or language; this binary distinction also seems exhaustive in the Nyāya context. While neither the NP nor the NM state the psychologistic criteria for a legitimate PA, such criteria are explicitly found in the *Nyāyabinduṭīkā*.[13] I describe it as follows: A PA is legitimate when it: 1) is "not contradicted" (*anirākrta*) by perception, inference (anumāna), general usage (*prasiddhi*) or one's own statement (*svavacana*);[14] and 3) generates an identical SV in the recipient's "mind". The latter's SV, by subsequent discursive communication, enables the recipient to then re-articulate a PA identical to, as so judged by the utterer, the original PA. Thus the SV ("sva", reflexive) is first person knowledge (sva ..., hereafter cited as 1st P), and the PA ("para" ... "other") is third person knowledge (cited hereafter as 3rd P). Hence 1st P anumāna is one of only two types of anumāna and 3rd P is the remaining type of anumāna.

I ask: if 1st P SV is translated as "inference", on what grounds could we justify the term "inference" rather than some other term, for noting the concomitance (vyāpti) relation of such cognitions as so judged by the trirūpahetu rule? The 3rd P PA exhibits some apparent features of a formalistic logic; however, it lacks formal deductive validity and the typical deductive procedure of matching an inference at issue with deductive inferential patterns such as modus ponens, $(((p \supset q) \,.\, p) \supset q)$, modus tolens $(((p \supset q) \,.\, {\sim} q) \supset {\sim}p)$ or (valid) transitive relations $(((A \supset B) \,.\, (B \supset C)) \supset (A \supset C))$. Hence there is somewhat more evidence for translating the 3rd P PA as "inference". But what of the 1st P SV?

My response regarding 1st P SV, is that there is very little evidence, in fact, almost no evidence at all, for translating the 1st P SV as "inference"; this is because the SV is 1st P, non-deductive and lacks the PA machinery.

5. THE TWO-VALUED CRITERION FOR A VALID DEDUCTIVE INFERENCE VS. THE POLY-VALUED CRITERIA FOR ANUMĀNA IN THE PA

There is another aspect of anumāna which when translated as "inference", leads the reader astray and constitutes an unrecognized dissimilarity between the PA and a true AE deductive

inference. It is this difference which is ignored, glossed over and buried when anumāna is translated as "inference" and the PA is translated as an "inference-for-others". To illustrate this aspect, I shall characterize the relevant concepts of an AE deductive inference and then contrast these aspects with the concept of anumāna in the PA.

While it is true that there is an ordinary language sense of "inference" which is loosely tied to coherent argumentation, and while this use of the word "inference" is legitimate, it is not a philosophical, metalogician's usage of "inference." The second use of this word is that of either a deductive or inductive inference; since there are none of the latter, our focus is the former.

When we are reading a nyāya logic text and we find the word "anumāna" translated as "inference", the reader naturally assumes the second sense of "inference", a precise, carefully structured <u>deductive</u> inference-schema. Furthermore, the reader is thus brought to believe that anumāna in PA must also be a relatively precise and explicitly structured <u>deductive</u> schema. The PA is emically structured, of course, except that there is <u>not</u> present any concept of a <u>deductively valid structure</u>.

A precise structure is, of course, exactly what a deductive inference is, and such an inference, when legitimate, has a precisely structured pattern or form. A desirable deductive inference, that is, a valid inference, is an inference which is of the same structure of one of (or is isomorphic with) a finite number of valid inference forms such as Modus Ponens, Modus Tolens or a valid Transitive Series. To illustrate, if we had the following sequence we could easily determine whether this was a valid inference.

> A: if <u>A</u> then <u>B</u>, and if <u>B</u> then <u>C</u>, and if <u>C</u> then not <u>D</u>, therefore, if <u>A</u> then <u>D</u>.

Of course, this is not a valid inference, for in the third conditional statement, the expression "if <u>C</u> then not <u>D</u>" violates the valid form. Consider the following expression.

> B: if <u>A</u> then <u>B</u>, and if <u>B</u> then <u>C</u>, and if <u>C</u> then <u>D</u>, therefore if <u>A</u> then <u>D</u>.

This inference (above, B) is a valid inference. Hence, to assert that portion of A "if <u>C</u> then not <u>D</u>" would be to make a mistake regarding the generation of a valid inference-schema. A general point can be drawn from this. With a relatively simple deductive inference pattern, if there is one single mistake in the form of the inference schema, that mistake is sufficient to render the inference schema an invalid one.

Also a specific point may be drawn from this. The criterion for an alleged AE inference schema being actually a desirable schema is whether it is isomorphic with a valid form. Thus in the example given above, because of the one mistake in the third conditional ("if C then not D"), compared with the potentially valid sequence of the remaining pattern of the inference, inference A is invalid. If the correction of A had been made, as above in B, the inference, a valid transitive series, would be valid. Thus the standards for judging a deductive inference to be valid admit of a strong disjunction, and no non-binary degree(s); the question of a deductive schema is whether it is an accurate assertion or denial that schema X is valid. By contrast the PA does admit a number of degrees in the standards for PA legitimacy.

The reason that the criterion for a legitimate PA admits of degrees is because of the emic psychologistic conception of the PA: this conception contrasts with the formal tradition of the AE deductive concepts of formal validity.

First, the sufficient condition for legitimate PA illustrates this psychologism. The sufficient condition for a PA, as I illustrated above, is when the speaker utters a PA which is isomorphic with his own SV such that the uttered PA generates a SV in the consciousness of the receiver which is identical with the original SV of the speaker. By the subsequent articulation of an isomorphic PA by the receiver, the comparisons of the two SVs and PAs may be made, and the legitimacy of the original PA may thus be established. An emphasis upon such psychological criteria would be abhorent to, and a mark of metalogical degeneration by, AE western logicians. However the Buddhist logicians continually emphasize the partial psychological nature of the PA.

Second, that the PA admits of controversial degrees of legitimacy is supported in the long and varied controversies in Indian philsophy regarding the required number of steps (avayava) (*members of the PA schema) in a PA. Sanghavi has noted the wide variety of views regarding the

> ... technique of presenting (prayoga-paripāṭī) ... a PA. (A variety of views in the different darśanas were presented.) Sāṅkhya logicians required three steps, namely the pratijñā (thesis), the hetu (justification) and the dṛṣṭānta (warrant).[15]

The logicians of the Nyāya Darśana admit and require five steps. Further, the Jaina logician, Hemacandra, held that the thesis and the justification "...are two minimum necessary steps while three, four, or five steps may be required in special cases."[16] As quoted by Sanghavi, Vādideva "... goes to the extent of granting that in dealing with a

particular type of hearers one single step, viz., the hetu might alone suffice, a position already granted by the Buddhist."[17]

Sanghavi summarizes in the following way:

> Thus "Māṇikyanadin" tells us that two steps and five steps are required in two different types of spheres, that is to say, two is the number of steps to be employed in the course of a debate (vāda-pradeśa) but either two or five steps may be employed (keeping in view the competence of the hearer concerned) in the course of a systematic exposition (śāstra-pradeśa). And what is to be remembered about Vādideva's stand is that he, like the Buddhist, grants that the hetu is the only step required when the hearer happens to be specifically learned (viśiṣṭa vidvān) (i.e. expert in the subject matter under consideration).[18]

Now let us make the contrasts explicit. The AE deductive inference is either valid or invalid; there are only two disjunctive possibilities. The PA-as-anumāna is a type of anumāna, both of which, I claim, have been misleadingly translated as "inference-for-others" and as "inference", respectively. The point at issue here is: what kind of "anumāna" is "parārthānumāna"?

First, I claim that the PA should not be translated as an "inference" because 1) it lacks the strict form of an AE inference; 2) it is not judged as a legitimate PA by the same deductive procedure as an AE inference which is judged to be valid if and only if the inference has a valid form; 3) if one presupposes that the PA is an inference, then it is most likely that one will not recognize 1) and 2); 4) one will assume that the PA is an implicit AE deductive inference; and 5) that the degree of distortion in the translation of a PA into an artificial language is acceptably small and that the artificial language translation is actually a fair, minimally distorting translation.

The PA admits of degrees of legitimacy as to its minimal requirements regarding its steps or components (the pakṣa, hetu, dṛṣṭānta, etc.); the true AE valid inference-schema does not admit of such degrees of legitimacy. Therefore the emic view of anumāna as in the PA, was significantly different. Because the translation of "anumāna" by "inference" obscures this distinction between emic degrees and the etic disjunctive AE valid or invalid choice, this translation "inference" distorts the expectations of the reader regarding the crucial distinction between the two. Hence this well-known translation is faulty and an alternative should be

found.

Thus the emic Buddhist logicians thought of anumāna in the PA as, first, a type of third person communicable ratiocination, mitigated by generic concepts (sāmānya-lakṣaṇa) superimposed upon the flashing flux of pratyakṣa; second, anumāna in the PA, admits of degrees of legitimacy which were contingent upon the presupposed knowledge of the recipient of the PA and the latter's intellectual nimbleness. "3rd person ratiocination" for "anumāna" as in the PA, is compatible with these assumed degrees of legitimacy; "inference" as in "inference-for-others" is not compatible with these degrees of legitimacy for an AE deductive inference does not admit of (such) degrees. An AE deductive inference is either valid or it is not valid; such a choice about a PA is not disjunctive.

An additional point should be noted. The quality of a PA admitting degrees of legitimacy contingent upon the communicative context of speaker and recipient, is completely compatible with the denial that the PA is isomorphic in form with a true AE deductive inference. It is the number of required members (avayava) of the PA which admits of degree, not the invariable emic form of the PA nor the metalogical rules or Nyāya metalogical theories of the PA. It is the latter emic PA form, rules and Nyāya theories which render the PA non-deductive and non-inferential; it is not the emic controversies over the allowable degrees of the required components which render the PA non-deductive and non-inferential.

6. ON ALTERNATIVE TRANSLATIONS OF PRATYAKṢA AND ANUMĀNA

Consider the following translations:

<u>Pratyakṣa</u>:

A) "Raw undifferentiated proto-perception."

<u>Anumāna</u>:

B) "1st P SV = 1st person differentiated ratiocination."

C) "3rd P PA = 3rd person differentiated ratiocination."

I draw the following conclusions about and justifications for these new translations. First, "pratyakṣa", by definition, cannot be what we usually mean by the ordinary language (=OL) term "perception". In OL, we usually assume that a perception exhibits some differentiation and presupposes the discrete isolatability of portions of what is perceived;[19] this cannot

be the case with Buddhist pratyakṣa, for such distinctions are, by definition, kalpanā and vikalpa, conceptual constructions, part of anumāna which is non-pratyakṣa. The words "raw" and "undifferentiated" convey this sense: "proto-perception" alludes to both that pratyakṣa may be the theoretical and epistemological basis for anumāna, and also that the subsequent process of knowing pratyakṣa via anumāna distinctions but knowing it as anumāna, has already occurred or has not yet occurred. Also "perception" in "proto-perception" conveys the idea of the theoretical basis of anumāna, but with the suffix "proto", the ordinary expectations of the possibility of any discriminated perceptions are thwarted; hence, a cautionary note is conveyed in this compound "proto-perception".

Second, let us consider "anumāna" as "ratiocination"; since 1st P SV is subjective (1st person) rather than inter-subjective third person knowledge, this point is conveyed in the translation "1st P" and "ratiocination." In other words, since "all our common, public knowledge," by Buddhist epistemological standards, is all anumāna, and that this knowledge is subdivided into "personal" and public cognitions via sāmānya-lakṣaṇa(s), then this important 1st P vs. 3rd P distinction is conveyed by "1st P SV". Next, "differentiated" conveys the discriminatory function of anumāna; "ratiocination" conveys the binary conceptualization of the concomitance criteria inherent in any anumāna cognition which meets the standards of the "three forms of the justifier" (trirūpa-hetu). Furthermore, "ratiocination" does not lead us to the illicit expectation that anumāna possesses the characteristic of any formalistic logic (which at best remains a subdivision of anumāna) such as formal validity or other deductive machinery.

Third, "3rd P PA" conveys these appropriate inter-subjective points noted above about "ratiocination," but avoids the misleading implications of anumāna-as-deductive-formal "inference"; also, "3rd P PA" does allow for the further analysis of whatever deductive features the PA may or may not possess without prejudging the formalistic characteristics of the PA by the illicit semantic influence of the very specialized term "inference". The older translation of the whole of "anumāna" as "inference" has, I believe, illicitly predisposed many Nyāya studies.

I now suggest that there are two areas or functions of anumāna only one of which is (almost) adequately conveyed by "inference". In short, "anumāna" carries a very great, varied epistemological burden and the translation "inference" is too narrow in its semantic range to carry this epistemological burden.

7. WHAT KIND OF ANUMĀNA IS THE PA?

The answer to the above question may be gleaned by examining the contrasting relations and expectations between the two questions which follow.

> 7.a. What kind of "inference" is "inference-for-the-sake-of-others"?
>
> 7.b. What kind of "anumāna" is "parārthānumāna"?

First, the PA is a particular sub-type of anumāna which is differentiated from other type of anumāna by the satisfaction of the implicit and explict metalogical rules, its structure and by means of its special vocabulary.

However, given my description and analysis of anumāna as pramāṇa, we find that the questionable translation of "anumāna" by "inference" leads the reader to assume that all these areas of cognitive discursive (anumāna) communication, first and third person, SV and PA, are somehow "inferential" because of the venerable translation "inference". In short, only one of three aspects of anumāna is somewhat similar to "inference" and 7.a. begs all the metalogical and epistemological questions.

What we discover when we look closely at the Buddhist texts, is that only one area of anumāna is (allegedly) inferential at all, namely, the PA; it lacks the PA form and is characterized as requiring only the satisfaction of the TRJ rule (trirūpahetu).

Elsewhere, I demonstrate that the PA is neither explicitly nor implicitly deductive, that it does not possess the metalogical theory and necessary conditions of the deductive concept of formal validity; nor does a PA justification exhibit the metalogical procedure of justification analogous to that of justifying a sound deductive argument. Hence the "inferentiality" of the PA is at least open to serious questioning. Thus there is some reasonable doubt regarding the exact "inferential" status of the PA. _A fortiori_, what of SV, which lacks most of the PA machinery?

Without the explicit requirement that the SV meet the PA form, rules and vocabulary, the SV does not come close to meeting even the questionable standards of the pseudo-deductive "inferentiality" supposedly possessed by the PA. However, my suggested translations of pratyakṣa and anumāna sweep away the misleading expectation that the two non-PA aspects of anumāna possess the crucial deductive concepts of AE validity, soundness, etc. The non-PA anumāna of the SV clearly requires the TRJ, but that requirement _alone_ does not

make it inferential.

To the contrary, the requirement that the SV meet the trirūpa-hetu criteria involves the examination of the threefold relation or ratios of the presence and absence, of an alleged concomitant binary set of properties. This examination of a possible concomitance is compatible with the translation of "anumāna" as "ratiocination", in the first person discursive vikalpa SV. The other descriptions in my justification of these new translations also apply.

8. SUMMARY

To summarize, the PA is a type of anumāna, but not necessarily a type of "inference". The PA is a formalistic type of third person, differentiated, discursive ratiocination possessing a constant, highly stylized form, quasi-variables, a metatheory of evaluating assertions, alleged concomitances and possible errors. And in this century, this PA has been alleged to be very similar to that of AE deductive logics; obviously, I strongly question this latter assumption.

The SV aspect of anumāna is distinguished by 1) its lack of the PA form and metalogical machinery (as noted in the preceding paragraph), 2) its first person locus and 3) its required satisfaction of the trirūpahetu criteria.

Having stated all the preceding, it is my opinion that the twentieth century habit of translating "anumāna" by "inference" has greatly misled and obscured the examination of all these aspects of anumāna. First, "inference" has misled the reader (non-Sanskritists and many modern Sanskritists too) in that it suggests that many true, deductive inferential features are also to be found in SV. Second, I suggest that 1) traditional aspects of anumāna (SV and PA) are a type of epistemological ratiocination about binary concomitances performed subjectively or in a public, highly stylized, structured way, rather than 2) an almost universalized process of proto-deductive inference making. Hence, 1) "anumāna" as a multi-sided type of ratiocination comes much closer to an accurate description and explanation of anumāna, as legitimate knowledge, a) anumāna as means and b) anumāna as cognition or knowledge (anumāna=jñāna).

This section started with the question: "What kind of anumāna is the PA?" The answer is that the Buddhist Nyāya PA is 1) a sub-type of Buddhist ratiocination which is 2) public, third person, differentiation, discursive, 3) in a highly stylized form, accompanied by, 4) a complex second order theory of 4) detecting concomitances to support a conclusion, for 5) detecting possible semantic and formalistic errors, by utilizing, 6) metalogical cliches (proto- or quasi-variables) and presumed structural relationships 7) for the general purpose of justification by offering procedural exemplars

(models) within a metalanguage; within the PA, the major emphases are the binary relations of properties and their exemplifications rather than AE validity, truth-preserving forms and transformational rules. Once again, there is no emphasis on or appeal to familiar valid inference patterns such as MP, MT, etc., nor any rules for truth-preserving (formal) transformation rules, such as transposition or De Morgan's theorem.

Thus the PA is a sub-type of Buddhist public ratiocination, not a sub-type of "inference". This point of translation has misled many and therein lies the question as well as the focus of subsequent inquiries.

Bowling Green State University

NOTES

1. Potter, K.H., ed.: 1977, *Indian Metaphysics and Epistemology: The Tradition of Nyāya-Vaiśesika up to Gaṅgeśa*, Princeton, Princeton University Press, pp. 184-185. This is the second volume of the *Encyclopedia of Indian Philosophies* of which Professor Potter is the general editor.

2. Bocheński, I.M.: 1961, *A History of Formal Logic*, Notre Dame University Press, p. 446, quoted in Potter, op. cit., p. 182.

3. In Potter's book, ibid., I have wondered whether on the bottom line of p. 182, a word seems missing; perhaps it is "free;" however, Professor Potter tells me that it should be read as it stands.

4. Tachikawa, M.: 1971, 'A Sixth-Century Manual of Indian Logic', in *Journal of Indian Philosophy*, 1, 128 and 144.

5. Tucci, G.: 1932, *The Nyāyamukha of Dignāga*, Heidelberg, p. 52; *Taishō*, 32, 3:C:2-9.

6. Sanghavi, S.: 1961, *Advanced Studies in Indian Logic and Metaphysics*, Indian Studies, Past and Present, Calcutta, pp. 74-75.

7. Matilal, B.K.: 1971, *Epistemology, Logic and Grammar in Indian Philosophical Analysis*, (Janua Linguarum), Mouton,

The Hague, pp. 34-35. For the many references to and studies of the author of NP, see Tachikawa, M.: 1971, 'A Sixth-Century Manual of Indian Logic', in Journal of Indian Philosophy 1, 119, note 3.

8. Ibid., p. 37.

9. Ibid., p. 38.

10. Hattori, M.: 1968, Dignāga on Perception, Harvard Oriental Series 47, Harvard University Press, Cambridge, p. 83.

11. Sanghavi, op. cit., pp. 77-78. Parentheses with an asterisk are the au+hor's additions and modifications.

12. Hattori, ibid., pp. 77-78.

13. Vinītadeva's Nyāyabindu-tīka, translated and annotated by M. Gangopadyaya, Indian Studies: Past and Present, 1971, Chapter Three, pp. 159-187, especially notes 1, 14, and 16. The number "1)" and "2)" and all parenthetical remarks preceded by an asterisk are the author's.

14. Ibid., p. 187.

15. Sanghavi, op. cit., 1961, 85-86; the author has modified this quotation.

16. Ibid., p. 86.

17. Ibid., p. 86.

18. Ibid., p. 87.

19. American College Dictionary, edited by C.L. Barnhart, New York, Harper and Brothers, 1953, p. 899: A. The isolatability of "a perception of something" is presupposed in the definition.

Mark Siderits

THE PRĀBHĀKARA MIMĀṂSĀ THEORY OF RELATED DESIGNATION

There are three well-known accounts of the cognition of sentence meaning in Indian philosophy of language: the pure sentence theory of the Grammarians, the designated relation theory of the Bhāṭṭa Mīmāṃsakas and the Naiyāyaikas, and the related designation theory of the Prābhākara Mīmāṃsakas.[1] The first and second of these theories are relatively well understood. Brough, for instance, has given a clear articulation and able defense of the sentence theory.[2] And the designated relation theory really requires little by way of introduction, since it is so close to what is probably the most widespread common-sense view of how we comprehend sentences--the view that we first grasp the meanings of individual words and then combine them to get a mutually related whole.

The related designation theory, on the other hand, does not appear to be as well understood as the other two theories. It is not immediately obvious what it means to say that the meaning of a word is the entity it designates in relation to the entities designated by other words occuring in a sentence in which that word is used. Nor has this theory received a clear and complete explication from modern scholars. Jha, for instance, gives a fairly accurate but all too brief and unilluminating sketch of the theory.[3] He also claims that the theory is supported by the Prābhākara position on language learning, which has it that the child can learn the meanings of words only from the use of injunctions. But as we shall see, all parties to the dispute can reconcile their respective theories of sentence comprehension with this claim about language learning. Staal[4] seems to think that one's position in the controversy between related designation and designated relation is determined by one's view on the question of whether sentence meanings should be seen as siddha (established states of affairs), or sādhya (actions which are enjoined to be done; roughly speaking this is equivalent to asking whether the statement-making sentence or the command is the more basic form of sentence). But in fact the dispute between related designation theorist and designated relation theorist transcends this issue as well. Staal is, I think, quite correct in his suggestion that the related designation theory is 'an extreme form of syncategorematicism,' but fails

B. K. Matilal and J. L. Shaw (eds.), Analytical Philosophy in Comparative Perspective, 253–297.

to make clear just what this means. Brough[5] gives the language learning story a different twist. He claims that related designation is supported by the view that we learn the meanings of individual words through the process of insertion and deletion (āvāpa and udvāpa). Yet here too we have an account of language learning which is perfectly acceptable to a designated relation theorist as well.

Bhishnupada Bhaṭṭacharya devotes a chapter of his A Study in Language and Meaning to the dispute between related designation and designated relation, and his introductory remarks[6] give a good summary of the respective positions. But his use of such terms as 'concept' and 'sense' throughout his discussion is quite misleading, and his detailed examination is largely confined to relatively late texts which present highly refined versions of the theories. Perhaps the clearest and most accurate discussion of the related designation theory is that of K. Kunjunni Raja. Through a detailed examination of the works of Śālikanātha, Kumārila, and others, he brings out some of the chief arguments for and objections against the theory. But he also repeats Jha's assertion that the theory follows from the claim that we learn language only from injunctions. And his summary comments are rather misleading, particularly his claim that the related designation theory is supported by 'the ubiquitous importance of context as a deciding factor in determining the meaning of a word.'[7] If I understand this correctly, he is here asserting that related designation is supported by such observations as that 'tiger' plays quite distinct roles in the two sentences, 'The zookeeper fed the tiger,' and 'The tiger is a carnivore,' referring in the first case to a particular tiger and in the second case to the species of tigers. But the Prābhākaras do not seem to have had such problems of disambiguation in mind when they designed their theory. They seem rather to have assumed that at least many of our words are strictly univocal in their literal uses, and they would claim that the issues their theory is meant to address arise even in a language with no ambiguous terms whatever.

I shall seek to present as clear and detailed a picture of the Prābhākara position as possible. My chief sources are Śālikanātha and Rāmānujācārya, but I shall also make use of Jayanta's discussion of related designation in Nyāyamañjarī. I shall examine some of the chief Prābhākara arguments for the theory, as well as the Prābhākara replies to some of the more common objections to the theory. What I ultimately hope to demonstrate is the importance of attaining a clear conception of the theory, not only because of the role it has played in Indian philosophy of language, but also because of the role it can play in current philosophical semantics.

As is often the case, it is useful to begin our investigation with a brief survey of Jayanta's comments on related designation. Although he ultimately rejects this theory, Jayanta presents an interesting overview of the position which better reveals some of its underlying motivation than do the writings of Śālikanātha and Rāmānujācārya. Thus in his prefatory remarks he indicates the centrality to the dispute of a question concerning the nature of language learning: 'Learning is most important; there is no ascertaining the meaning of speech without learning. But is learning to the sentence meaning from the sentence, or to the word meaning from the word? If learning is to sentence meaning from sentence, that is related designation; if to word meaning from word, designated relation.'[8] If one holds that words can be learned in isolation, then one will naturally be attracted to the view that the meaning of a word is the independent entity designated by that word. If, however, one holds that such learning is impossible, that only complete sentences provide sufficient data to the language learner, then it might seem more plausible to suppose that the meaning of a word is an entity in relation to those entities designated by the other words in a sentence in which it occurs.

The Prābhākara goes on to argue for the claim that language learning takes place only with respect to complete sentences, not individual words. First he points out that words are not used in isolation.[9] Now it is an obvious truism that the sentence is the basic unit of communication, but we might be inclined to wonder about the relevance of this observation for questions about language learning. Thus we might believe we acquire at least the rudiments of our language by means of ostension, and in the typical case of learning by ostension a word is used in isolation. But the Prābhākara reveals the full thrust of his argument when he notes that 'there is no worldly conduct [performed] by means of a word.'[10] For the Prābhākara, as for most Indian philosophers of language, language use is always action-oriented: the purpose of an utterance is always to affect, directly or indirectly, the behavior of the audience. It is clear that such purposes are not achieved through the utterance of isolated words. (Seeming counter-examples, such as 'the door', are best thought of as elliptical sentences.) But this fact about language use places special constraints on our theory of language learning. To learn a language is to learn to use utterances to affect the conduct of others; thus the learning situation must involve observations of competent speakers uttering and obeying commands. This is not to say that ostension can play no part whatever in language learning. But when we consider the typical case of ostensive definition in the light of the present point about the purposive nature of linguistic activity, we can see that its role is at best minor. Suppose that I

have learned to utter 'red' when a red object is pointed to. We might say that I have then learned to obey a certain (implicit) command. But I am not yet able so much as to turn the tables and command others to name the color, let alone issue or obey commands concerning the fetching of red objects. The mastery of all this requires observation of more typical language use, use involving complete sentences.

Śālikanātha and Rāmānujācārya have far more to say on the topic of language learning and the lessons to be drawn from it for semantics; we shall later consider in some detail their own arguments. We shall then see that Jayanta is wrong in supposing that the related designation theory follows from the claim that language learning must initially be from sentence to sentence meaning. For not only is this claim compatible with a pure sentence theory, but when it is supplemented by the insertion-deletion account of word mastery it becomes compatible with designated relation as well. Jayanta's discussion is useful, however, in that it indicates the background assumptions which make the topic of language learning relevant to the question of word meaning and sentence meaning. A second point on which Jayanta's treatment of related designation is similarly significant has to do with the semantic contribution of a word to the sentence in which it occurs. Here he provides two striking analogies for the competing views of word meaning, neither of which is employed by our two Prābhākara authors. The first comes in his statement of the general Prābhākara criticism of designated relation: 'That view which would have it that what are denoted by words are word meanings which are related to one another like a row of stakes is also not right, since there is no such kind of worldly behavior, and because of the difficulty of making out relation afterwards, since we do not perceive the operation [of a word] after it has ceased [i.e., performed its function of designating its meaning].'[11] The analogy of a row of stakes expresses well the Prābhākara complaint that if the words of a sentence do nothing more than designate unrelated entities, it is difficult to see how we arrive at the related aggregate meaning which serves the purpose of conduct. Of course the designated relation theorist has several ways of trying to meet this criticism, and these will be examined below. But it is useful to think of the general difficulty being faced here as analogous to the task of showing how a row of stakes can be turned into some sort of functional whole.

The second and more significant analogy is introduced in the following:

> In the case of an aggregate, it is not found that the aggregated elements each perform their own operations individually [separate from those of the others]. Just as in the case of wagon

> parts we never say, 'There is this part, which is made from this, there is that [part] which is made from that; these wagon parts, functioning separately, perform the operation of a wagon.' Just so, a word is never used alone. Even where it is used, there is nothing which can be taken as that part of the effect which is its own. But from its operation together with other words, a word effects just a related meaning.[12]

It is also employed in a subsequent argument to the effect that on the Prābhākara view words may be said to have meanings. 'It is not the case that the meaning of a word cannot then be known, as is seen when we divide up the operations of these [words] through the method of insertion and deletion, as with the wagon parts.'[13]

Thus on the related designation theory a sentence is to be viewed not as an indivisible whole, but as a functional aggregate or complex like a wagon. Now what is striking about this analogy is that we are tempted to think of a wagon as made up of parts which perform independent functions. This temptation stems, no doubt, from the fact that the parts of a wagon are themselves discrete substances which can be physically isolated by disassembling the wagon. But it is wrong to suppose that the functional contribution of a wagon-part to the wagon as a whole is likewise discrete or independent. No amount of investigation of an isolated axle can ever reveal to us the role it plays in the functioning of a wagon. The role of the axle is just to facilitate the turning of the wheels and thus the passage of the wagon over the ground. This function can only be performed in conjunction with the functioning of wheels, yoke, bed, etc. And this function cannot be discovered except by comparing the behavior of the wagon with the axle attached to its behavior when the axle is removed.

Here too there seems to be room for dispute, for one might object that the manner in which one learns the meaning of a word--by noting how its presence and absence in sentence frames affect conduct--tells us nothing about the nature of word meaning in general. On this point Jayanta is also silent. The value of his discussion, once again, lies in the overview it gives us of the theory of related designation. For detailed argumentation we must look to our Prābhākara authors, and so we now turn to an examination of Vākyārthmātṛkā and Tantrarahasya.

It is helpful to think of the Prābhākara view as an attempt to find a middle ground between a pure sentence theory and the designated relation theory. This helps explain the efforts Śālikanātha and Rāmānujācārya expend in arguing against a pure sentence theory. The theorist who holds a middle position is apt to be mistaken by an opponent from one

extreme as holding the position of the opposite extreme. As Jayanta has pointed out, the Prābhākara holds that the sentence is the most fundamental unit of meaning. To a word theorist like the Bhāṭṭa Mīmāṃsā or the Naiyāyika, this may sound like the view of Bhartṛhari, that words are not intrinsically meaningful but are rather just useful fictions produced by grammatical analysis, like verbal roots and phonemes. But the Prābhākara insists that a sentence is made up of real words, that words do play a semantic role in the comprehension of sentences, and that words may thus be thought of as having meanings. It is to establish this difference between his own position and that of the pure sentence theory that the Prābhākara argues against the latter.

Śālikanātha and Rāmānujācārya both give the following argument against the sentence theory. It is agreed that the child begins to learn language by noting how the conduct of its elders is affected by the utterance of sentences. According to the sentence theory, the child thus learns that a given sentence has the capacity to designate (i.e., means) a certain action. According to the related designation theory, the child does not stop here but goes on to discover, through the process of insertion and deletion, that the words of a sentence have designative powers. (The sentence theorist holds that only sentences have designative powers, the word theorist that only words have such power.) Now consider the following eight sentences: 'Śiśa, bring the cow.' 'Śiśa, tie the cow.' 'Vatsa, bring the cow.' 'Vatsa, tie the cow.' 'Arbhaka, bring the cow.' 'Arbhaka, tie the cow.' 'Ḍimbha, bring the cow.' 'Ḍimbha, tie the cow.' For the sentence theorist, the mastery of these eight sentences will require the child to learn eight distinct designative powers; for the word theorist, however, since these sentences are composed of seven words this requires the mastery of just seven powers. (The definite article is not employed in the Sanskrit equivalent.) Moreover we can produce eight more sentences by inserting 'white' before 'cow'. For the sentence theorist this will mean a further eight powers, but for the word theorist just one. Thus a word theory is to be preferred on the grounds of parsimony.[14]

Now arguments from economy tend to be viewed with distrust, in large part because economies often prove to be false: the lightening effected by shedding one type of posit in a theory is frequently more than offset by the compensating adjustments which must be made elsewhere in the theory. Here the argument is perfectly legitimate, however. In structure it is similar to Chomsky's well-known 'infinite capacity, finite resources' argument, although it does not make use of the questionable assumption that a competent speaker of the language is able to understand infinitely many sentences. Instead we are simply given a fragment of the Sanskrit corpus

and shown how the child can come to understand these sixteen sentences by learning the designative powers of eight words. Notice it is not claimed that the child cannot master any of these sentences until he has learned the meanings of the component words. As we shall see in more detail below, the Prābhākara maintains that at the initial stage in the learning process it is only whole sentences whose meanings are mastered. He merely wants us to consider the consequence of supposing that language learning must always proceed in this fashion. The result is a veritable explosion in the number of distinct sentence meanings which must be mastered one by one. It is clearly preferable to suppose that the child, after mastering some basic stock of sentences, learns to exploit the phonetic, syntactic, and semantic regularities exhibited in that stock in order to achieve mastery of additional sentences.

Rāmānujācārya advances a particularly telling version of this argument when he points out that on the sentence theory a sentence and its negation can bear no discernable semantic relation to one another: the denial of a given sentence transforms it into a distinct sentence of indeterminate meaning.[15] Suppose a child to have mastered some 100 commands of the 'Devadatta, fetch the cow' variety. She then learns the corresponding prohibitions for four of the original commands. Surely it is odd to suppose that she must learn the meanings of each of the remaining prohibitions individually. For surely she must be able, at some stage in the learning process, to recognize the signficance of the negative particle 'mā' in prohibitions. There is an important datum to be explained here, namely the facility with which the chld masters sizeable portions of the corpus once the language learning process is under way. A sentence theory is unable to explain this fact, while a word theory gives a simple and more plausible explanation.

The novel sentences objection[16] is another variant of this basic theme. It is unlikely that the reader has previously encountered the sentence, 'Devadatta the cowherd eats three bowls of rice a day.' Yet we understand it. How can this be accounted for on anything like a sentence theory? Since on that theory the sentence is an indivisible semantic unit, and since we understand the meaning of a sentence when we know its satisfaction conditions, our ability to understand this sentence could be explained only by supposing that we had prior knowledge of the relation between it and its satisfaction conditions. An adequate explanation of this ability requires acceptance of the composition principle, the principle that the meaning of a sentence is a function of the meanings of the words out of which it is made up.

The Prābhākara thus rejects the sentence theory as incompatible with the composition principle. But the

intuition which underlies the sentence theory--that words are not used in isolation--should not be neglected. To ignore the significance of this insight is to run the risk of falling into the opposite extreme, the designated relation theory. Exclusive reliance on the composition principle leads one to suppose that words have independent meanings, and that one comprehends a sentence by first apprehending the isolated meanings designated by the individual words of the sentence and then somehow putting those meanings in the proper relations with one another. The Prābhākara agrees with the designated relation theorist that sentence meaning is a function of word meaning. He takes issue, however, with the claim that words have independent designative capacities, for he also agrees with the sentence theorist that one should not overlook the significance of the fact that the sentence is the basic unit of language use. But how is the significance of this fact to be acknowledged within the confines of a word theory? The Prābhākara answer is that our account of word meaning must be constructed in accordance with a second principle as well, the context principle: only in the context of a sentence does a word have meaning.

The semantic theory which emerges when both the context principle and the composition principle are given due weight is the related designation theory. According to this theory, the meaning of a word in a sentence is the entity designated by that word in relation to the entities designated by the other words in the sentence. A central issue in the dispute between the Prābhākara and the designated relation theorist concerns the scope of the designative capacity of words. On the designated relation view, this capacity stops with an independent entity: the designative or referential relation is between a word in any occurrence and a determinate entity. According to the related designation view, the designative capacity of a word does not stop with an entity but includes that entity's relations with other entities as well: the designation relation is between a word in some sentential context and a determinate entity with its relations to other entities.

Notice that on this view as I have just formulated it, a word cannot be said to have a designative capacity if it occurs apart from some sentential environment. Frege's views on the sense of a predicate expression are in some ways like the Prābhākara position on word meaning. Frege held that the senses of such expressions are incomplete, and thus naturally demand completion by the senses provided by other words in a sentence. But he also held that the referents of such expressions are likewise incomplete. This resulted in the posit of unsaturated objects, and Frege has been extensively criticized for this seemingly odd addition to our ontology. It is not a consequence of the Prābhākara position that there are unsaturated objects. This possibility is blocked by

the stipulation that a word's designative capacity is exercised only within a sentential context. It is not clear whether the Prābhākaras were motivated on this point at least in part by the desire to avoid ontologically suspect entities. In any event, the result is a commitment to the context principle which is stronger than Frege's. Since the meaning of a word is its designatum, and a word designates only with the help of other words, a word cannot be said to have meaning in isolation but only in the context of a sentence.

If, as the Prābhākaras have argued, words must be said to have meanings, how can it be shown that the designative power of a word does not stop at a determinate entity but proceeds to include its relations to other entities? Our authors provide one basic argument against the general position of designated relation, and then go on to give objections to the two distinct formulations of that position. The general argument concerns the nature of language learning, and thus involves implicit acceptance of the maxim that we discover what word meaning is by looking to see what we learn when we learn the meaning of a word. The basic account of language learning is accepted by all the parties to the dispute: The child masters such commands as 'Bring the cow,' 'Tie the horse,' etc., by repeatedly observing one elder issue such commands and another behave in accordance with them. It is noteworthy that the stock of learning sentences consists exclusively of injunctions. Although this is never explicitly argued for, the assumption here seems to be that the satisfaction conditions of an injunction--some purposive human behavior--will prove more perspicuous than the truth conditions of a statement.

At this stage the child has learned that the sentence 'Bring the cow' designates the action, the bringing of a cow. At the next stage, the child performs the operations of insertion and deletion on the stock of learned sentences. Insertion of 'cow' into the sentence frame 'Bring the ___' results in a sentence whose satisfaction conditions are the bringing of a cow. Deletion of 'cow' from this frame (that is, its replacement by another word) results in a sentence whose satisfaction conditions are the bringing of something other than a cow. The meaning of 'cow' should then be the difference between the first and second sets of satisfaction conditions. (This result must of course be verified by looking at the behavior of 'cow' in other sentence frames.)

In this way the child learns that words possess a distinctive property, the power to designate meanings.[17] It is here that the dispute begins. We might be tempted to suppose that the difference obtained by performing the methods of insertion and deletion on 'cow' is just a cow.[18] It is this temptation which leads to the designated relation theory.

And, argues Śālikanātha, it should be resisted, since 'insertion and deletion are in accordance with relation, i.e., not exceeding relation. Thus in "Bring a cow" there is insertion of cow only as related to bringing, and in "Bring a buffalo" there is deletion [of cow] only as related to that [bringing].'[19] Or as Rāmānujācārya says, 'Insertion and deletion have related domains.'[20] His argument is that one learns of the designative power of words only by observing the conduct of the elders which occurs upon the utterance of an injunction. One infers a cognition in the addressed elder which was produced by the utterance and led in turn to the conduct. And one further infers a power in the utterance to produce this cognition, namely designative power. But when insertion and deletion are performed upon a set of utterances, and the designative power of the utterance is apportioned among its constitutent words, we must bear in mind that the designated action is made up of entities in relation to one another. There is no bringing which is not the bringing of some animal or other, and there is no cow which is not the object of some action or other. It is thus illegitimate to suppose that the operations of insertion and deletion yield an unrelated cow as the designatum of 'cow', for that would involve going significantly beyond the data. All that we are justified in concluding is that 'cow' when inserted in the sentence frame 'Bring the ___' designates a cow related to bringing, when inserted in the frame 'Tie the ___' designates a cow related to tying, and the like.

Here the designated relation theorist objects that while conduct does indeed have related entities as its object, it does not follow that words must be related designators. For it is possible that words designate unrelated entities, and that an intermediate operation, namely awareness of unrelated word meanings, brings about cognition of the relations among those entities.[21] Śālikanātha replies as follows:

> Passing over words with their primary relations, it is not right to base the power to make sentence meaning known in meanings. It is beyond dispute that words are the designators. It is thereby agreed to that they have the power to designate, it can be readily supposed that it [the power] includes relation. But there is posited the power to produce comprehension in word meanings instead. 'In terms of lightness, better to posit a property than a property-possessor,' the power of related designation is properly posited just of words.[22]

Here we have a straightforward argument from economy. The claim is that it is more parsimonious to suppose that the

ability to make relation known resides in the generally acknowledged designative capacity of words, than it is to posit a separate power for making known relation whose locus is the designated word meanings. The latter supposition appears to be forced on the designated relation theorist since, as Jayanta has pointed out, it seems odd to say that a word which has already exercised its power--designating an independent entity--can then make a new effort to indicate relation. The opponent is thus stuck with the claim that a set of stakes is endowed with the capacity to bring about cognition of their relatedness. And, argues Śālikanātha, why posit this extra power when we can so easily endow the designative power of words with the property of bringing about apprehension of relation?

It was remarked earlier that arguments from economy should be approached with caution. This is a case in point. It remains to be seen what price must be paid for endowing the designative power of words with this property, and how on the alternative proposal word meanings can come to have such a power. The first task is best accomplished by looking at some of the standard objections to related designation and the Prābhākara replies. The second task will require us to investigate the two distinct formulations of designated relation offered by the Bhāṭṭa and the Naiyāyikas respectively. Only when all of this is done will we be in a position to assess the claim that related designation is the more parsimonious of the two theories.

There is, however, another objection to designated relation that is sufficiently general that we can examine it before beginning our investigation of the two versions of the theory. Kumārila had claimed that it is not as odd as it might seem to endow word meanings with the power to bring about apprehension of mutual relation, since we see that entities have such powers in cases other than verbal cognition. Thus when we indistinctly perceive a patch of white, hear the sound of neighing, and hear the striking of hooves, we apprehend that a white horse is running, that is, whitness, hoofbeats, and neighing are mutually related through the substance horse.[23] The cognitive processes which lead to this conclusion are described by the Bhāṭṭa as follows. One infers the universal horseness from this sound of neighing, and one infers the action of galloping from the rapidity of repetition of the hoofbeats. One then apprehends that the whiteness cannot inhere in any substance other than a horse, like a cow, by realizing the incompatibility of a cow being where horseness is. Finally one cognizes by means of presumption (arthāpatti) the mutual relation of white, horseness, and galloping as fit things requiring relation to another entity, in the form, 'A white horse is galloping.' It is crucial to this account that one becomes aware of the general relatedness of

white, horseness, and galloping solely from the force of the expectancy, proximity, and fitness of these entities; presumption is effective here only in supplying the particular horse through which these are mutually related. This is crucial because only thus will the example serve to show that objects cognized by means other than speech themselves bring about awareness of relation.

The Prābhākara response to all this is that in such cases there is no awareness of relation without the use of inference and presumption. In particular it is insisted that one could not have the cognition, 'A white horse is galloping,' unless one were already aware that the locus of whiteness is the same as that of horseness and galloping. One knows that whiteness inheres in a horse by means of presumption: from the inferred presence of a horse where the neighing is perceived and the rapid motion is inferred, one concludes that the substance in which whiteness inheres could not be other than a horse. And this presumption could not be carried out unless one already knew that the place where the white patch was seen is the same as the place where neighing and hoofbeats were heard. Otherwise one would merely infer the existence of some substance in which white inheres; the relation of this substance to the substance which is the ground of neighing and galloping could not be known.[24] Thus the perceived objects are not themselves responsible for cognition of relation. There is, then, no known case outside the area of dispute where entities which are cognized as unrelated make us aware of relation; and by the rules of debate this means that designated relation is unproven.

Nor is this all. As this example shows, the designated relation theorist holds that it is just the apprehension of word meanings that brings about cognition of mutual relation. In that case it is no longer clear that speech is the means of knowledge whereby sentence meaning is known. The pramāṇa of speech is defined by the opponent as the knowledge of relation caused by speech by means of an intermediate operation of designating word meanings; and this is clearly not the same thing as the knowledge of relation caused by apprehension of word meanings. Now it might be objected that the latter has as its object word meanings which are apprehended by means of speech. But this does not show that speech is the pramāṇa here, for the validity (prāmāṇya) of the apprehension of relation belongs to that alone and not to the cognition of word meanings from which it proceeded.[25] The resultant divorce of speech from cognition of relation has dire consequences. The designated relation theorist is no longer able to rule out the possibility that sentence meaning will include relation to entities cognized by means other than speech. Thus when Devadatta is told to bring the cow, he is justified in bringing the cow and the goat, since he cognizes by perception that the

cow is related to the goat by physical proximity.[26]

There is an obvious way out of this difficulty for the designated relation theorist. He need merely say that words, in addition to designating unrelated meanings, somehow endow the latter with the ability to make known relation. In this case the cognition of relation is brought about not by mere apprehension of word meanings but by speech itself through an intermediate operation. Thus one insures the verbalness of sentence meaning and guards against the possibility of relation to entities not designated by words. There are two views on how this occurs. On the first view, words possess yet another power in addition to the power to designate word meanings, the power to assign to word meanings the power to designate relation. The second view eschews the posit of any additional power to designate relation, as violating parsimony; instead it holds that designated unrelated word meanings lead to comprehension of relation through secondary implication (laksanā). This is the position of the Bhāṭṭas. We shall examine each view in turn.

While both Śālikanātha and Rāmānujācārya describe the view that posits a power-assigning power,[27] only Rāmānujācārya ascribes this view to the Nyāya, and that only by implication. The accepted position of the school is that the relational element is manifested by sentence-meaning (vākyārthatā).[28] Rāmānujācārya describes the school as holding that 'the meaning of a sentence is the recalled own-meanings which are designated by words, or the word meanings which are supplied with expectancy, etc.'[29] Here there is no mention of an additional power of words, yet Rāmānujācārya immediately goes on to charge the Naiyāyika with the lack of parsimony in posited powers. There must, he feels, be some factor which is responsible for our not stopping with a mere list of entities but instead going on to apprehend the entities in mutual relation. Rāmānujācārya reasons as follows:

> On the designation thesis the power of words to bring about comprehension of sentence meaning is the power of words each to designate its own meaning. Then since one does not experience the making known of sentence meaning in the case of entities made known by pramāṇas other than speech, and one does not experience this in the case of what is stated, words add the power to bring about comprehension of sentence meaning to word meanings. Their power to effect [comprehension] is posterior.[30]

That is, the Naiyāyika position must involve the posit of a power-assigning power, since otherwise we are unable to explain the difference between speech and the cognition of

isolated entities through other pramānas.

It is problematic what form such a power might take, given the overall Nyāya view of meaning. One possibility, which was taken up by Jayanta, is that words possess a purpose-revealing power (tātparyaśakti). In the initial language-learning situation we discover that speech is uttered for the purpose of bringing about cognition of related entities. We also learn, through the processes of insertion and deletion, that each word has the power to designate an independent entity. Thus the competent speaker, upon hearing a sentence, will cognize the set of distinct entities designated by the words, but will also, by virtue of his recognition of the power of words to serve this general purpose of speakers, realize that the distinct designata stand in need of mutual relation. This leads to cognition of word meanings related in accordance with the expectancy, proximity, and fitness of the words.

This view is actually quite close to that of the Bhāṭṭa. It differs from the later chiefly in its posit of an actual power; Kumārila seeks to avoid what he sees as lack of parsimony by using implication instead as the device which takes us from word meanings to sentence meaning. The chief Prābhākara objection to this position is that it involves the posit of three powers--the designative power of words, the power-assigning power of words, and the relation-comprehending power of word meanings--and is thus less economical than a theory which gives words the power to designate meanings in mutual relation. We shall defer consideration of this parsimony argument. Another line of attack focuses on the supposition that we employ the purposefulness of speech in obtaining knowledge of sentence meaning. Since this sort of criticism is also used as an objection to the Bhāṭṭa thesis of implication, we shall likewise defer consideration of it until we have examined that thesis.

A more orthodox Nyāya approach than Jayanta's does not explicitly posit a power-assigning power in words. Instead it claims that the force of expectancy, proximity, and fitness leads from word meaning to sentence meaning. This may be what is meant by the statement that relation is manifested by 'sentence-meaningness'. It is generally agreed that in order for a string of words to constitute a sentence, they must possess mutual expectancy, proximity, and fitness. Thus as one apprehends the entities designated by the words, one is also aware that each word 'requires' other words in the string, that these words have been uttered close to one another in time, and that there is no obstacle to the relation of the designated entities. These factors will then lead one to the apprehension of the entities in mutual relation. This view is more chary than that which posits a power of purpose, but it is still really tantamount to a power-assigning power theory.

It achieves its economy of powers by locating the origin of that force which takes us from word meaning to mutual relation in those properties of words which have already been acknowledged to be necessary conditions of sentencehood. In effect the claim is that these three factors endow word meanings with the power to bring about cognition of relation.

Of the three, it is expectancy which is the most important in this task. And it is on expectancy that the Prābhākara criticism of this position focuses. Expectancy is, roughly, the demand produced by one word for another word to complete its sense. When we hear 'bring' we naturally expect some additional word to be uttered; and if none is forthcoming, we are at a loss as to what to do. It is this phenomenon which the notion of expectancy is meant to explain. The Naiyāyikas define expectancy as the inability of a word to give rise to a presentational apprehension (anubhava) without the use of other words. For this they have been criticized by the Mīmāṃsakas on the grounds that expectancy is a kind of desire, and words cannot be said to have desires. The Mīmāṃsaka insists that expectancy be defined as the human desire to know the remainder. But this dispute seems to be over terminology alone. Both sides agree that we experience a desire to know the remainder when we hear 'bring'. And they likewise agree that this desire must have some basis in the linguistic phenomena themselves. The Mīmāṃsaka reserves the term 'expectancy' for the desire itself, while the Naiyạ eyika applies it to the cause of this desire. The important issue, though, concerns the nature of the source of this desire.

The need for some objective basis of the expectancy of words should be clear. It is commonly said that a complete sentence is one which expresses a complete thought. And presumably a thought is complete when we no longer require further information. As Śālikanātha points out, however, this makes the notion of a complete sentence relative to the hearer's desires. It is apparent that on hearing 'bring' one will desire to know the noun in the accusative case which is to be supplied. But 'the end of the desire to know is not ascertained. For just as there can be such a desire to know the case, so there can occur the desire to know that he [the agent] is such and such a person--of that color--performing the action, desire to know of things other than that case, etc. Or else, because of lack of purpose things other than that case are not desired to be known.'[31] Unless we posit some objective basis of the desire to know, 'Bring the cow' will sometimes count as a complete sentence, sometimes not.

What, then, is the source of expectancy? On the Nyāya view under consideration, it is a necessary connection: 'Thus a verb possesses necessary connecton with a noun; upon hearing that [verb] there would be the desire to know the noun. And

similarly with a noun.'[32] This cannot be right, however. As Rāmānujācārya points out, 'There are infinitely many word meanings which possess necessary connection,'[33] and so once again the end of the desire to know cannot be ascertained. This view would have it that the verb produces the desire to know the accusative-case noun because an action possesses a necessary connection with an object. But an action likewise possesses a necessary connection with an instrument, which in turn possesses a necessary connecton with its material cause, and so on. The upshot is that 'Bring the cow' is <u>never</u> a complete sentence.

In fact the only satisfactory explanation of expectancy is that which employs the supposition that words are related designators. Śālikanātha's solution to the problem we have been considering is this: 'Incompleteness of designation and incompleteness of designatum are the basis of the desire to know the remainder.'[34] That is, we wish to know the remainder when we hear 'bring' because it is an objective feature of the designative function of the word and of the meaning designated by the word that they are both incomplete. With respect to the designative function this incompleteness stems from the fact that the word designates an entity in relation to those entities designated by other words in the sentence. Thus the designative function cannot be completed until those other entities are also designated. 'It is not possible that there is related designation in a word having an unuttered proximate correlate.'[35] With respect to the designatum this incompleteness consists in the fact that we have been given a relation without its correlate, that is, something of the form aR__. The incompleteness of the designatum is shown by the fact that it comes with a hole which requires filling.

The claim which is being made here is an important one for semantics in general. The claim is that we cannot give an adequate account of sentencehood unless we suppose that words are related designators. We obviously require some way of determining which strings of words are to count as complete sentences. The notion of expectancy is clearly useful here. For it permits us to say that a string is a complete sentence just in case the expectancy produced by each word in the string is satisfied by some other word or words in the string. And it seems right to make expectancy a psychological property, for we do experience a sense of anticipation when a sentence is interrupted prior to completion. It is crucial, however, that this psychological state be given some objective grounding, since otherwise the notion of expectancy cannot perform its allotted task: we cannot use it to explain the notion of a complete sentence if the upshot is that the degree of inquisitiveness of the hearer determines whether a sentence is complete or not. We might, like the Naiyāyika, base our sense of expectancy on necessary connections among designata.

This, however, leaves us with the imposing task of determining just which necessary connections are essential to sentencehood and which are not; as we have seen, the necessary connection between an action and its instrument need not give rise to an expectancy for an instrumental upon hearing a verb. Or we might suppose that it is syntactical properties which are the source of expectancy. The child learns that conduct follows upon the utterance of a word with an imperative ending only when it is followed by the utterance of a word with an accusative ending. She thus learns that if the one ending occurs, the sentence is not complete unless the other ending also occurs. Now it is true that this sort of syntactical approach, which explains our sense of expectancy by way of the associations that are developed between sets of inflectional endings, might yield a solution to the problem of finding a criterion of sentencehood. Thus it might at least in principle prove possible to construct a purely formal grammar of the language, a procedure for determining, for any string of words, whether it constitutes a sentence or not from its formal properties alone. But the Indian philosopher of language has rather a deeper question in mind. He seeks not merely a criterion of sentencehood, but rather an account of the notion of a complete sentence which explains the fact that sentences yield cognition of related meanings while individual words do not. And this the syntax proposal as so far presented does not seem to do.

We might remedy this, however, by extending the syntax proposal in the following way. Words designate unrelated entities. Syntactical relations among words indicate the relations among those entities. On hearing 'bring' we understand that the action of bringing is designated, but we also expect some other word or words to follow because we have learned that a transitive verb in the imperative only occurs when accompanied by a noun in the accusative. On hearing 'cow' we understand that a cow has been designated. We then go on to cognize the relation between bringing and cow because of our awareness that the accusative of 'cow' satisfies the expectancy aroused by 'bring'. There is no distinct element in the sentence which designates relation among word meanings; there are just the formal relations among the words--and these are enough to take us from word meaning to sentence meaning.

One still wants to know just how this is done, however. It seems as if syntactical relations among words could motivate us to place word meanings in mutual relations only if we already somehow knew that syntax is in general intended to serve this function. But then we have reverted to the thesis that we get sentence meaning from word meanings by way of an implication which makes use of the purposiveness of speech. Otherwise we find ourselves stuck with a string of unrelated word meanings plus an additional datum--an indicator of the

relations among these meanings--which we have no incentive to employ. The Prābhākara line of argument is that this dilemma faces anyone who supposes that words designate independent entities. The only way to avoid it is to hold that words designate entities in relation to other entities. It is quite natural that we should have a sense of expectancy when we hear 'bring', since we do not yet know what the word designates, and we wish in general to understand the meaning of an utterance. We understand the meaning of 'bring' only when we subsequently hear 'cow', for then we know that 'bring' here designates bringing in relation to a cow. We can abandon our search for the source of the glue which will hold word meanings together in sentence meaning, for the designative function of a word is such that word meanings never stand in the unrelated state.

This, then, is the Prābhākara response to those like the Naiyāyikas who would make sentential properties such as expectancy responsible for apprehension of relation. We turn now to the Bhāṭṭa proposal, which has it that sentence meaning is implied by word meanings. The official position of the school is described as follows. Word meanings bring about comprehension of sentence meaning by means of secondary implication (<u>lakṣaṇā</u>). And secondary implication occurs because one is aware, having cognized the unrelated entities designated by the words in a sentence, that one has not apprehended what is intended to be expressed. For we discovered when we first learned language that speech is uttered for the purpose of bringing about comprehension of a single related meaning.[36] That is, we go from word meanings to sentence meaning by way of our realization that cognition of a set of unrelated entities does not serve the general purposes of speech.

Our Prābhākara authors begin their examination of this position with a discussion of secondary implication. The term 'lakṣaṇā' is often translated as 'metaphor', since many of the cases included under this head involve metaphorical transfer of meaning. A standard example is, 'The village is on the Ganges'. Here 'on the Ganges' (<u>gaṅgāyām</u>) must be taken in a metaphorical or extended sense, since if taken literally it would mean that the village is located in the river, and this we know cannot be. This has led some theorists to suppose that a necessary condition of lakṣaṇā is that the literal meaning of the term be incompatible with the meanings of the other words in the sentence. It is then objected that the Bhāṭṭa account of sentence meaning does not involve lakṣaṇā, since there is, for instance, no incompatibility between the meanings of 'bring' and 'cow'. The Bhāṭṭa response is that lakṣaṇā involves more than just what we would call metaphor (which at least often involves literal falsity), but synecdoche as well (e.g., 'The armada was made up of fifty sail'). The relevant difference here is that when

we compute the meaning of a metaphor we exclude the literal meaning of the term in question ('in the river') and replace it with a related meaning, whereas in working out the meaning of a synecdoche we do not exclude the literal meaning of the term but rather add to that one or more related meanings. Thus in our example, 'sail' stands for the whole, ship, of which its literal meaning is a part; here the literal meaning is not excluded but extended.[37] There can thus be lakṣaṇā in the cognition of sentence meaning insofar as the meaning of 'bring' (a simple bringing) is not excluded but extended to include relation to a cow when we go from word meanings to sentence meaning.

Śālikanātha's response to this is that the Bhāṭṭa has incorrectly defined lakṣaṇā. On the Prābhākara view lakṣaṇā requires both non-apprehension of relation, and connection: we perceive that the literal meaning of the term in question does not allow of relation to the meanings of the other terms in the sentence, and the desire to cognize relation leads us to find something connected to the literal meaning of the term which does allow of such relation. 'Bring the cow' satisfies neither condition. Bringing is not unfit for relation with cow, and since bringing is related to cow, it cannot be that some meaning connected to bringing is also related to cow.[38] This is hardly a telling objection, however. In fact the whole dispute so far seems quite muddled. What the Bhāṭṭa would seem to want to say is that apprehension of sentence meaning is like the standard case of metaphor in that it involves seeming incompatibility. But here the incompatibility is not between word meanings but rather between the cognition of a set of unrelated entities on the one hand, and the knowledge that speech is employed with the intention of bringing about cognition of a single related whole on the other hand. Awareness of this incompatibility leads one to reject the set of unrelated entities in favor of the same entities in mutual relation. It is true that on this account the mechanism whereby one computes sentence meaning from word meanings can be called lakṣaṇā only in a sense which is itself metaphorical. But pointing this out does not show the view to be unworkable.

Rāmānujācārya has a more interesting objection to this proposal. 'Besides, it is unproven that sentence meaning is itself implied, because cognition is all at once, just as in word meaning, without consideration of non-apprehension of connection.'[39] That is, we are not aware of such an intermediate operation of considering the incongruity of a set of unrelated entities when we compute the meaning of a sentence, and so this account must be false. I have argued elsewhere[40] that such appeals to our intuitions concerning linguistic cognition can play at best a very small role in semantics. But this appeal does not seem entirely out of place. In the case of a

normal metaphor (as opposed to 'frozen' metaphors, where the extended sense has taken on the status of a convention) we seem to be aware of the incongruity which results from placing the literal meaning of the metaphorically occurring term in that environment. We exhibit such awareness for instance by the manner in which we express appreciation of a well-turned figure of speech. We frequently laugh, and this not just in admiration but also to release the tension produced by the search for the appropriate non-literal meaning. And this search in turn can only occur if we are aware of the oddness which results from taking the word literally. (Needless to say, we do not always go through this process; often there are other cues, such as intonation and context, which tell us that a word is not being used in its ordinary sense.) The fact that it often takes time to work out the meaning of a metaphor shows our awareness of the literal incongruity of the expression. In the case of a sentence all of whose words are to be taken literally, however, the process of comprehension would seem to be quite different. We are certainly never aware of the oddness of getting a list when what we wanted is a single mutually related whole. Nor does the competent speaker seem to need additional time, after hearing the words of a sentence and apprehending their meanings, to compute the meaning of a sentence. When we reflect on those processes of linguistic comprehension of which we are aware, I suspect we will agree that apprehending the meaning of a sentence is not at all like working out the meaning of a metaphor in the way that the Bhāṭṭa thesis claims it is.

The trouble with this sort of appeal to our intuitions is that we can never be sure we are aware of all the processes involved in even the most basic mental operations. The concept of nirvikalpa perception is a case in point. There is, however, an underlying line of argument to this objection which is rather more compelling. In essence the objection is that there is a crucial disanalogy between cognition of sentence meaning and cognition of metaphorical meaning. Now Rāmānujācārya locates the source of the disanalogy in our awareness of incongruity. It is, I think, correct to focus on the element of incongruity which is central to metaphor, though perhaps not on our awareness of it. What is important here is that by its nature metaphor must involve some incompatibility between literal meaning and environment. This means that the metaphorical use of a term must always be exceptional. When such usage comes to be expected by the speakers of a language we have a frozen metaphor, which soon becomes an additional literal meaning of the term. It could hardly be the case that the comprehension of sentence meaning from word meaning is exceptional. As a rule we use words to convey sentence meaning. Given our awareness of this fact, we should expect the development of something akin to the

phenomenon of frozen metaphor. That is, the competent speaker of a language should come to expect that words will in general be used to express their own meanings in relation to other meanings. But it is difficult to see how this differs from saying that words have the capacity to designate related meanings. What makes it so odd to say that sentence meaning is comprehended in a manner analogous to comprehension of metaphor is that in the case of sentence meaning the intermediate operation is so readily bypassed. Because the ultimate aim of word usage is generally just comprehension of related meaning, we should quickly drop the assumption that words designate unrelated meanings--the source of the alleged incongruity--and assume instead that words designate meanings in relation to other word meanings.

This line of argument applies equally to the Bhāṭṭa lakṣaṇā thesis and to theories like Jayanta's which posit an intention-manifesting power in words as a way of more directly producing the alleged incongruity. Śālikanātha gives an argument which is not as general in scope. It is directed against the view that apprehension of relation is achieved by means of the consideration that the utterance of speech would be pointless if it did not lead to cognition of relation, since then it would not serve the purpose of conduct. This view explicitly denies that there is any power, such as the power to manifest intention, that functions in the apprehension of relation. Presumably this is done for reasons of economy. But, argues Śālikanātha, in this case Vedic sentences would be meaningless. We have no means of knowing the meanings of Vedic sentences apart from those sentences themselves. Thus we have no grounds for asserting that such sentences must lead to cognition of relation in order to serve a purpose. Moreover, it seems odd to say that in general we should be struck by the purposelessness of perceiving a set of unrelated entities, for we seem able to pass quite readily from cognition of activity to its relation to a goal, without the intermediate step of reflecting that this activity is pointless without relation to the goal.

Suppose, however, that we grant the awareness of pointlessness. This then leads to cognition of relation by way of our knowledge of the general purpose of speech, which was acquired at the time of language learning. Now presumably we gained this knowledge by observing the conduct of the elders which followed upon the utterance of injunctions. 'The being intended, of words, for relation of what is expectant, fit, and proximate is apprehended in the conduct of the elders.'[41] But how, Śālikanātha asks, is this possible if neither words nor word meanings have the ability to bring about awareness of relation? The conduct we observe consists of related entities--the cow in relation to the act of bringing. Words by assumption bring about cognition of unrelated entities.

The argument is that we could come to know that speech is intended to produce cognition of relation only if we already knew that something produced by speech--either words or cognition of word meanings--had the capacity to bring about awareness of relation. Otherwise we should simply be at a loss to explain how the gap between word meaning and conduct was closed. In general we attribute an intention to an agent only if we believe the agent to believe that his or her actions have the capacity to bring about the desired goal. The case of speech should be no different.

This leaves us with two possibilities: that words make known sentence meaning, and that word meanings make known sentence meaning. But, argues Śālikanātha, since it is agreed that words have the power to designate word meanings, it is only natural to suppose that they have the power to designate relation as well.[42] Such views as Jayanta's are to be rejected not only because they are unparsimonious (that too is argued here), but also because it seems odd to place the power to bring about cognition of sentence meaning in something other than words. This is a recurring theme in the Prābhākara criticism of designated relation--that the latter theory makes linguistic cognition non-verbal. The Prābhākara notion seems to be that we should use words to do as much of the work as possible in sentence comprehension. This stems no doubt at least in part from the fact that they wish to maintain both that speech is a distinct pramāṇa and that the Vedas are authorless. The first assumption will motivate us to seek out distinctly verbal causes for all the aspects of linguistic comprehension. The second assumption rules out recourse to speaker's intention as a way of explaining various facets of the phenomenon. But even if we reject both assumptions, their admonition still seems sensible. It sounds decidedly odd to say that while words convey their individual meanings, something else conveys the relation of those meanings in sentence meaning. For the most part, the processing of sentence meaning proceeds in a simple and straightforward manner. Since we obviously do make use of the words in arriving at their own-meanings, this suggests that we rely on them as well as for comprehension of relation; bringing in consideration of the designated entities simply seems to be an unwarranted complication of the matter. Granted, this can hardly be decisive in settling the dispute between related designation and designated relation. It does, however, help account for some of the plausibility of the former thesis.

We have considered some of the chief Prābhākara objections to the designated relation theory. We now turn to the principal objections to the related designation theory. As we shall see, in answering these objections our authors introduce several refinements to their position. One question we should keep in mind is whether these refinements ultimately undercut

the seeming plausibility which the theory gets from its simplicity.

Of the various objections to the related designation theory, that of mutual locus is the most frequently cited and perhaps the most troublesome. It arises out of a dilemma which is formulated by Śālikanātha as follows:

> [Assume that] a related own-meaning is being designated by a word--is it designated as related to another word meaning which is [already] designated, or not [yet] designated? It must be one or the other. If not, then the use of other words is pointless. And so it follows that from one [word] there is cognition of all relations. If it is [designated as] related to what is designated [by another word], then that word, by virtue of its related designatingness, depends for its designation on the meaning obtained from another word, hence there obtains [the fallacy of] mutual locus.[43]

If in the sentence 'Bring the cow', 'bring' is able by itself to designate bringing in relation to a cow, then on hearing the first word we should comprehend the meaning of the sentence, so that the utterance of 'cow' is unnecessary. This is obviously false, since we seldom grasp the meaning of a complete sentence from just the first word. Thus it must be that 'bring' designates bringing related to a cow only after 'cow' has designated its meaning. But on the related designation thesis, the meaning of 'cow' in this sentence is a cow in relation to bringing. Thus 'cow' cannot succeed in designating its meaning until 'bring' has exercised its designative function. We seem to be caught in a kind of 'hermeneutic circle': for each word in a sentence, we cannot comprehend its meaning until we have comprehended the meanings of the other words. This is the fault of mutual locus.

In Ṛjuvimalapañcikā, Śālikanātha summarizes his response to this objection as follows:

> There is no ascertainment whatever of meaning from a single word alone; however, having heard as many words as there are in a sentence, having ascertained the qualifiers of number and case in accordance with rules, the meaning is apprehended. There, the words having been heard first, it is just related word meanings which are recalled, just as on seeing a word meaning the word [is recalled]. Even if the related is being designated, own-form does not disappear, by this it is possible for a word to recall the

> own-form of that by means of proximity. Thus a word designates own-meaning related to word meaning brought into proximity by memory, there is no mutual locus. For there is no dependence in memory, recalled word meanings themselves bring about apprehension of mutual relation.[44]

Here we must first explain what is meant by the 'own-form' of a designatum. It is obvious that the various literal uses of 'cow' all have at least this much in common, that they designate a cow in relation to other things. Now the related designation theorist denies that a cow is the meaning or designatum of 'cow', but this is not to deny that this one type of substance, the cow, is invariably present in every case of designation by the word. The cow is then the own-form of the meaning of 'cow'. This is also expressed by saying that the cow is the own-meaning of 'cow.'[45]

Now when we learn a language we acquire <u>samskāras</u> (memory traces) which are activated whenever we hear a learned word. Śālikanātha is here claiming that these own-meanings or core meanings are recalled by the saṃskāras. This might sound as if he were flirting with the notion that word meanings are unrelated entities, but in fact he takes great pains to block this suggestion. First, this is a case of mere recollection, which obviously plays a role in linguistic cognition but should not be mistaken for a case of cognition itself. Second, the core meaning is not all that is recalled by a word. Śālikanātha stresses this when he says, 'Speech produces the recollection of just the own-form <u>as well</u>.'[46] The recalling of own-form is said to be 'included within' the grasping of connection which is the apprehension of sentence meaning. In fact, what is recalled upon hearing a word is not just its core meaning, but that core meaning as it occurred within the learning sentences, i.e., in its related state. Since, however, the operations of insertion and deletion require a variety of sentence frames, the core meaning will be recollected as related to a large variety of different entities. To learn the meaning of 'cow' we must encounter it not only in 'Bring the cow', but in 'Tie the cow', 'Milk the cow', etc. Clearly we will not recall all of the many entities to which the core meaning was related in previously cognized sentence meanings involving the word. We will, however, recall the core meaning's relatedness, the fact that as it occurred in the meanings of learning sentences it was related to other word meanings.

This last point is crucial, for there is a natural tendency to suppose that in learning the meaning of a word we drop the relational element by a process of abstraction. That is, we suppose that the core meaning of 'cow' must be all that we are left with once we have seen the diversity of entities

with which it can enter into relation in sentence meanings---bringing, tying, milking, white, etc. Indeed Rāmānujācārya sometimes suggests that this is the picture he has of word meaning, for instance when he says, 'Immediately upon the articulation of a word there arises the recollecting of an unrelated word meaning.'[47] This impression is undercut, however, by his reply to the objection that related designators could not serve to recall own-meanings: 'If the meaning is related, then the recalling of that [own-meaning] occurs even for the being known of the unrelated part, but of the correlate with indistinctness, that is not an object of memory.'[48] By 'unrelated part' and 'correlate' he means that as yet unascertained entity which is related to the meaning of the word being heard. If 'Bring the elephant' is a novel sentence, then we obviously cannot, upon hearing 'bring', recall bringing as related to an elephant. But we can recall that 'bring' has always been used to designate bringing in relation to some kind of substance. The elephant and its relation to bringing are not recalled, they are not objects of memory at this stage of sentence comprehension; but they are 'known with indistinctness', that is, known under the description, 'a related substance of the appropriate sort.'

It is important to bear in mind here that the Prābhākara does not admit the existence of 'incomplete' or 'unsaturated' objects. While what is recalled on hearing 'bring' might be depicted as having the form 'bringing-R-__', this is not to say that there is such an object as bringing-R-__. A (binary) relation exists only when both of its relata are given. This is, once again, one of the points of the Prābhākara allegiance to the context principle. If a word can have meaning only in the context of a sentence, then the related designation theory is not faced with the consequence that word meanings are unsaturated objects.

The first stage of sentence comprehension is then the recalling for each word in the sentence of its core meaning and the fact of its relatedness. As Śālikanātha puts it, 'He whose connection grasping saṃskāra is uncorrupted, having heard the word remembers this: this is expressive of what is related to an expectant, proximate, and semantically fit correlate. Thus by the remembering of just what is recollected, the own-form, even though [as yet] unrelated, is of the things which participate in relation.'[49] The second stage of sentence comprehension involves the application of rules to the recalled but as yet unrelated own-meanings. In the Ṛjuvimalapañcikā passage which was quoted above, Śālikanātha mentions only rules pertaining to case and number. Rāmānujācārya provides a longer list, however:

> Subsequently there occurs in the hearer a consideration, concerning these heard things, having as object the various own meanings. [This consideration] is of the nature: this is a single sentence, that is distinct sentences; this is the intended meaning, that is not the intended meaning; this is metaphorical, that is literal; this is capable of relation, this is principal, that is subsidiary; this is to be enjoined, that is not to be enjoined, and the like.[50]

Evidently this stage of processing involves the arranging of the recalled own meanings into a unified sentence meaning through the use of the available syntactic cues and a set of semantic rules. This is clearly thought of as a complex process, for the rules have different weights assigned to them. Thus for instance the rule of single sentencehood, to the effect that whenever possible a string of words should be considered a single sentence, has precedence over the rule that whenever possible words are to be taken in their literal senses. For we are told that metaphor and secondary usage are resorted to in sentence comprehension precisely in order to avoid 'division of sentences', that is turning a string into two or more sentences due to incompatibility of some of the literal meanings.[51]

The most basic rules for sentence comprehension, however, are those pertaining to determination of the principal and subsidiary words in a sentence. These play a crucial role because they determine the order in which expectancy is generated. Suppose, for instance, that we have recalled the core meanings of each of the three words in the sentence, 'Bring the cow with-a-stick.' Here 'bring' is the principal word, and so our desire to know its correlate occurs first. We then consider the remaining words and find that 'cow' is a semantically fit correlate whose own expectancy for relation to an action is satisfied by relation to bringing. The process might stop here were it not for the fact that the rule of single sentencehood requires us to try to make a single sentence out of all the words in the string. Thus we return to consider 'with-a-stick', which generates a desire to know its correlate, namely a semantically fit action, and find that bringing is just such an action. Thus we arrive at comprehension of the sentence meaning, the bringing of a cow by means of a stick.[52]

What we have here, then, is a set of rules determining the notion of a minimal sentence and its structure, along with rules determining the structure of more complex sentences.[53] These rules are thought of as governing the order and manner

in which we compute the meaning of a sentence. And it is by following these that we arrive at a unified sentence meaning. In the final stage of sentence comprehension, the 'state of designation', 'all those words which are brought under the dominion of recollection are simultaneously causes of awareness. There being relation by means of the secondary-primary relation, from the oneness of the primary [word] there is oneness of sentence meaning.'[54] By following the order in which expectancy is generated and satisfied, beginning with the minimal sentence and proceeding outward, we have attained awareness of the relations among the words. Now, all the words being presented simultaneously by an act of memory, we achieve synthesis of their recalled own-meanings in a mutually related sentence meaning by means of the ascertained relations. This is called the state of designation because it is only at this point that the words of the sentence can be said to have achieved the status of designators. When we intially encounter a word in a word string, 'the word which is being heard is not a related designator because its correlate is not yet proximate, and because related designation is dependent on that.'[55] Only when the related correlate of its core meaning has been brought to awareness can we say that a word has discharged its designative function. As Śālikanātha has already said, 'A word designates own meaning related to word meaning brought into proximity by memory.'[56] 'Bring' succeeds in designating only when we are able to hold its recalled own-meaning, bringing, in relation to the recalled own-meanings of 'cow' and 'with-a-stick' in a single act of apprehension.

The Prābhākara theory answers the mutual locus objection by distinguishing between recollection of a word's core meaning, which occurs immediately on hearing the word, and the cognition of the word's designatum, its own-meaning in relation to the meanings of the other words in the sentence, which occurs only after all the words have been heard and their roles worked out in accordnce with the rules governing order of expectancy. The mutual locus objection was based on the assumption that the semantic contribution of a word to a sentence should be revealed all at once upon hearing the word. And this is clearly impossible if words are related designators. Now this assumption makes sense if our basic view of sentence comprehension is like that of the designated relation theorist. If we suppose that we get completed word meanings one-by-one as the utterance of the sentence proceeds, then somehow put these all into mutual relation afterwards, we might wonder how the first word in a string could possibly bring about cognition of its core meaning in relation to the core meaning of a word which is as yet unuttered. And seeing that this is impossible, we might then conclude that the word could make no semantic contribution whatever. But this is

question-begging, for it is only the designated relation theorist who believes that a word is capable of designating its meaning in isolation, without the assistance of other words. It is true that the word must initially reveal some part of its meaning if the processing of the sentence is to get off the ground. On the Prābhākara account it reveals, through recollection, its core meaning together with the fact of its relatedness. But the fact that the own-meaning is thus revealed as related to an as yet unknown correlate shows the words's designative function to be as yet incomplete, and so it generates expectancy. And similarly with the other words in the string. When these expectancies have all been satisfied in the appropriate order, the own-meanings are assembled in a collective act of memory and placed in their proper relations. It is now that the complete semantic contribution of the word to the sentence is made known.

Now it might be objected that this introduction of a stage in which as yet unrelated own-meanings are all assembled in memory and placed in mutual relation makes the Prābhākara account too cumbersome. Śālikanātha points out, however, that the use of an intermediate stage of collective memory is unavoidable on the designated relation theory as well. 'On that view also what are to be designated by all words are unrelated own meanings. It is to be affirmed that apprehension of sentence meaning comes after all these have been brought to memory.'[57] If each word brings about cognition of its meaning as it is heard, then in order to apprehend the meanings in mutual relation they must first be gathered together in memory after the utterance of the sentence is completed. Thus the related designation theory cannot be charged with lack of parsimony on this account, since the stage of collective memory must be resorted to in any event.

A second common objection to the related designation theory is the infinite correlates objection. This is, as usual, stated with succinctness and clarity in Ṛjuvimalapañcikā:

> Moreover, there is heaviness of posited powers if related designation is agreed to. Is it designated as related to the universal, or to the particular? On the first hypothesis, this designation is useless, since that is established as well by the fitness (sāmarthya) of the meaning. But on the second, there is no possibility of grasping connection, by the infinity of particulars. There being expressiveness of an ungrasped relation in what was heard before as well, the result is cognition of any meaning whatever.[58]

By 'related to the universal' is meant the relation which the core meaning of a word bears to the universal inhering in all the fit correlates of the entity. To say, for instance, that 'bring' designates bringing as related to the universal is to say that bringing is designated as related to the universal of substanceness. If this is the sense in which a word is a related designator, however, then the inclusion of relation in the designative function of a word seems superfluous, since this much of the relatedness of the own-meaning can be known from its expectancy for a semantically fit correlate. It is also objected[59] that sentences give rise to apprehension of relation with the particular. Since it is claimed by the related designation theorist that the words are the source of all that we cognize in sentence meaning, relation with the particular correlate of the core meaning should also be designated by a word. But this leads to the infinite correlates objection proper: 'The objects of verbal designative power that would have to be posited are infinite, being the objects of comprehension of own-meanings which are related to infinitely many correlates.'[60] Since the action of bringing may stand in relation to any of the infinitely many things inhered in by substanceness, the word 'bring' must be said to have infinitely many designative powers, one for each of these possible relational complexes. The result is that it becomes impossible to grasp the designative power of the word, that is, learn its meaning.

Śālikanātha responds by likening words to vision. It is agreed that vision has but one power, the power of seeing. Yet vision is capable of bringing about a cognition of color only as inhering in various particular correlates--a particular pot, a particular cow, etc. The fact that vision is able to reveal color as related to infinitely many correlates is not taken to show that vision must possess infinitely many manifesting powers. By the same token, it is argued, we should not take the fact that 'cow' may be used to designate a cow in relation to one or another particular act of bringing, tying, milking, etc., to show that the word possesses infinitely many designative powers. Like vision it has but a single power, which will manifest itself differently depending on the correlate which is brought into proximity by the other words of a sentence in which it is used.[61]

This response may appear unsatisfactory in that it seems to conflate an epistemological issue with one that is ontological. The problem was that it is unclear how we could ever learn the meaning of, e.g., 'cow', if that meaning is a cow in relation to infinitely many correlates. Assuming that vision does have just one manifesting power, it is far from evident how the fact that that power is one in spite of its revealing one thing (the universal of colorness) in relation to infinitely many particulars solves the problem. What is at issue,

it might be objected, is not whether the power of the word is one or many, but rather how we can achieve mastery of a word whose designative power is in any event clearly multifarious.

Such suspicions are misplaced, however, for Śālikanātha's response is germane. The relevance of the analogy with vision is to be found in the notion of a common core. It is true that the cognitive content of the vast majority of vision episodes is completely disparate. In three successive moments I see black inhering in a blackboard, green inhering in a desk, and then black again, but this time inhering in a print. There is something common, however, namely cognition of color inhering in a particular color which in turn inheres in a particular substance. That the particular colors and the substances in which they inhere are infinite in number does not alter the fact that we can isolate this common element by abstraction. When it comes to characterizing the object of visual cognition, we have a choice. We might describe it as one universal, colorness, possessing an infinite number of relations of the general form '-inherence$_n$-particular color$_n$-inherence$_n$-particular substance$_n$'. Or we might describe it as an infinite set of relational complexes of the general form 'colorness$_n$-inherence$_n$-particular color$_n$-inherence$_n$-particular substance$_n$'. Now from the perspective of a realistic view of relations, which is the common view of Indian metaphysics, our first characterization of this object seems decidedly unpalatable. It is highly implausible to suppose that there can be an entity which simultaneously stands in infinitely many relations to infinitely many particulars.[62] It seems more likely that the object of an episode of vision is one member of the infinite set of relational complexes. Here our ontological preferences are entirely beside the point, however, since it remains true in either case that we can discover a common core which is made manifest by all instances of vision. To say that vision has a single manifesting power is merely to say that we can find some way of characterizing such a common core; it is not necessarily to commit ourselves to the existence of some such entity as the general object of vision.

By the same token we can say that 'cow' has a single determinate designative power, namely the power to designate a cow in relation to whatever semantically fit correlate is brought into proximity by the other words of a sentence in which it occurs. The fact that there are infinitely many such correlates is no barrier to our effectively characterizing the object of this power. It is true that we might have qualms about positing the existence of a core meaning which possesses infinitely many relations to infinitely many particulars. And these qualms might lead us to say instead that the word has an infinite set of designata. This in turn could lead us to conclude that the word possesses infinitely many designative powers, and thus that its meaning cannot be mastered. But the

last conclusion is not warranted, for the ontological question is irrelevant. If the Prābhākara held that a word can exercise its designative power in isolation, then we might have to decide whether there is a single power whose object is a core meaning related to infinitely many correlates, or infinitely many powers mapped onto an infinite set of relational complexes. As we have seen, however, they hold that a word exercises its designative power only within the context of a sentence. For each occurrence of the word, there is a determinate relational complex which is its designatum. All we need do, then, to determine whether or not the word has a single power and thus can be mastered, is see if we can effectively characterize what is common to all the members of the set of designated relational complexes. And this we already know we can do.

There is a sense in which the Prābhākara denies that there is such a thing as _the_ meaning of 'cow'. The opponent who raises the infinite correlates objection sees that this is a consequence of the related designation theory, but has failed to appreciate the sense in which it is asserted. Under a reference theory of meaning, the meaning of a term would be whatever the term is generally used to designate. According to the related designation theory, a word is generally used to designate some one entity in relation to those other entities designated by the other words in a sentence in which it occurs. Since any term is generally used in a wide variety of different sentential environments, there will be no common designatum for all a word's literal uses. This does not mean that we cannot master its use, however. We have already seen how this is done. Through observation of a sufficent stock of learning sentences in which it occurs, together with observation of the satisfaction conditions of those sentences, we come to recognize that 'cow' designates a cow in relation to bringing, tying, etc. When we subsequently encounter the novel sentence, 'Feed the cow', hearing 'cow' causes us to recall the word's own-meaning, as well as its having been used to designate this own-meaning in relation to such actions as bringing and tying. This recollection, together with a similarly structured recollection concerning the use of 'feed', together enable us to arrive at the designatum of 'cow' in this instance, namely a cow in relation to feeding. As Rāmānujācārya puts it, at the moment of hearing 'cow' the relation of cow to feeding is discerned 'indistinctly'; that is, we know from our recollection of past uses of the word that it is here being used to designate a cow in relation to some fit correlate, but do not yet distinctly cognize the correlate. That cognition will come only after we hear 'feed'. But then through collective memory and a rule-governed process of synthesis we achieve comprehension of sentence meaning. That the designatum of 'cow' in this sentence is quite new to us,

that we have never before seen the word used to designate cow in relation to feeding, is no bar to our understanding the meaning of the sentence. And clearly, nothing could better show our mastery of a word than our ability to understand novel sentences in which it occurs.

Jayanta records another objection to the related designation theory which is not mentioned by either of our Prābhākara authors.[63] This concerns the processing of such semantically deviant sentences as, 'One hundred herds of elephants are standing on a fingertip.' It is assumed that this sentence is literally meaningless, since the designated objects are not fit to be related in the manner indicated in the sentence. The objection is that on the related designation view the sentence must nevertheless succeed in placing these entities in mutual relation. The argument seems to be that we can know these entities are not fit for relation only when the words of the sentence have completed their designative functions; and these functions are completed only when they bring about cognition of mutual relation. The Bhāṭṭa opponent points out that on the designated relation theory there is no ascertainment of relation, since sentence meaning is arrived at by means of the expectancy, proximity, and semantic fitness of the word meanings. That is, on this account we first cognize the unrelated entities designated by the words of the sentence, then become aware that some of these entities are not fit to stand in the relations which the expectancies of the words would require. This awareness of unfitness prevents us from going on to apprehend the entities in mutual relation.

At this point it is claimed that such an expedient is not open to the Prābhākara, and thus that he has no satisfactory explanation of the semantic deviance of this sentence. If the related designation theorist claims that semantic unfitness blocks awareness of relation, then he is asserting that unfitness blocks designation, which seems _prima facie_ absurd: how can we know that x and y are not fit for relation when x and y have not yet been designated? Now Śālikanātha suggested one way out of this difficulty when he stated that in the stage of recollection a word recalls its core meaning as related to what is expectant, fit, and proximate. The related designation theorist might perfectly well say that in the 'stage of consideration' we become aware that the recalled own-meanings of the words are not fit for relation. But this is not the solution which Jayanta reports the Prābhākara as adopting. Instead he is described as embracing the conclusion that semantically deviant sentences give rise to awareness of relation. We compute the sentence meaning in the usual fashion and then become aware that this meaning is contradicted by knowledge obtained from another pramāna. This leads us to reject the computed sentence meaning as semantically deviant. In the case of the present example, we already

know (perhaps from verbal testimony, perhaps by means of inference) that even a single elephant cannot stand on a fingertip. The surface meaning of the sentence is thus contradicted by a well-entrenched item of knowledge, and thus the sentence must be considered literally meaningless. This account of semantic deviance is said to recommend itself by virtue of its congruence with the Mīmāṃsā doctrine of the intrinsic validity and extrinsic invalidity of presentative cognitions (anubhava). The apprehension produced by a sentence is to be excluded from the class of valid cognitions only if, subsequent to our awareness of that apprehension, we realize that it is contradicted by a cognition which is known to be valid. The doctrine of intrinsic validity might be summarized as, 'A presentative apprehension is to be presumed valid until such time as it is proven invalid.' Clearly a semantically deviant sentence cannot be presumed to be a valid means of knowledge if we are unable to apprehend its meaning.[64]

A second argument in favor of the claim that semantically deviant sentences do bring about awareness of relation is more interesting. When we hear the sentence, 'A hundred herds of elephants....,' we grasp the indication of substratum, superstratum, and action.

> If, however, this sentence were devoid of relation, it would indicate no more than does the string of phonemes ka ca ta ta pa etc. Here as well related designation, having as it does the syntax which we find in such [non-deviant sentences] as, 'There are ten pomegranates,' is not prevented. That which contradicts is just the object of another [pramāṇa], it is not the object of [the pramāṇa of] verbal relation.[65]

Here the tables have been quite neatly turned on the opponent, for the designated relation theorist is now stuck with the task of explaining how, on his theory, a semantically deviant sentence can bring about an episode of awareness which is quite unlike that produced by a string of nonsense syllables. According to that theory, awareness of the mutual relatedness of word meanings is brought about through expectancy, proximity, and fitness. Since the last of these requirements is not fulfilled, there should be no awareness of sentence meaning. We are clearly aware, however, of what would be asserted by this sentence if it were possible for it to be true. In effect the designated relation theorist wishes to incorporate a set of selection restrictions into the computation of sentence meaning. In processing a semantically deviant sentence, he wants to say, we first get the unrelated word meanings from the words, then begin to assemble these into a related whole

in accordance with expectancy and proximity, but are stopped by a selection restriction, derived from our global knowlecge of how entities stand to one another in the world, which tells us that two of the designated entities are not capable of mutual relation. But this does not explain our awareness that the sentence places incompatible entities in mutual relation. Perhaps we might interpose a stage in which, having assembled the unrelated meanings in accordance with expectancy and proximity, we achieve a sort of tentative awareness of mutual relation; this is followed by a check for violations of selection restrictions, and if none are found we go on to achieve full cognition of sentence meaning. This seems extraordinarily cumbersome, however. It would be far simpler and far more plausible to say that a word in any event designates its own-meaning in relation to the own-meanings of the other words in the sentence. A semantically deviant sentence is simply a special case of a sentence which we find to be false.

According to Jayanta, the related designation theorist holds that a semantically deviant sentence is a sentence whose meaning is contradicted by a distinct valid cognition. This has the unfortunate consequence that the sentence 'Devadatta is in the house,' becomes semantically deviant if we have just seen Devadatta leave the house. For the sentence designates Devadatta and the house in the contained-container relation, and we know by means of another <u>pramāna</u> that these entities do not stand in this relation. The same result follows for the designated relation theorist, however, for the trouble stems from the definition of semantic fitness. If semantic fitness is simply defined as lack of contradiction by another <u>pramāna</u>, then any sentence which we know to be false will turn out to be semantically deviant. Thus some define semantic fitness as the capacity (<u>sāmarthya</u>) for relation, or as the possession of a form which is conducive to relation (<u>anvayaprayojaka</u>). But these notions are unclear, and they become especially murky when they are coupled with the epistemic requirement that the capacity or form be known by some pramāṇa.[66] A better definition of semantic fitness is that it is the absence of being qualified by a valid cognition whose correlate is the constant absence of relation between the meaning in question and that with which it is asserted to be related.[67] This keeps 'Devadatta is in the house' from turning out to be semantically deviant, since there is no valid cognition of the constant absence of relation between Devadatta and the house. But what about 'Pi is a rational number'? Surely this sentence is not semantically deviant, yet we have the best possible evidence for the constant absence of inherence of rationality in pi. This difficulty will infect the Prābhākara and the designated relation theorist alike. My point, however, is that I am not sure there is a way of drawing a precise boundary between semantically

deviant sentences and sentences which are well-formed but false. The sentences that we all agree are semantically deviant would seem to be just the ones which are contradicted by well-entrenched and highly confirmed items of global knowledge. Since entrenchment and globalness come in degrees, it is to be expected that the semantically deviant should shade off into the merely false. But if this is right, then the related designation theory has a decided advantage over the designated relation theory in accounting for the phenomenon of semantic deviance. With a sentence which is simply false, we recognize its falsity by first ascertaining its meaning and subsequently realizing that it is contradicted by a valid means of knowledge. If semantically deviant sentences are just a special case of false sentences, then an account of the processing of semantically deviant sentences which follows this basic pattern is clearly preferable to one that has us stop the processing before we achieve full cognition of sentence meaning.

This concludes our examination of some principal objections to the related designation theory. It will be recalled that we deferred consideration of several parsimony arguments until we were in a better position to assess the overall merits of the competing theories. It is now time to redeem this pledge. The two arguments in question were these: (1) against the general view that word meanings possess the power to make mutual relation known, that it is more parsimonious to posit a property (namely the property of making relation known, which is to be a property of the power of words to designate meanings) than to posit a property-possessor (namely the said power of word meanings, which possesses the property of making relation known); and (2) against the view that words possess a power-assigning power which endows word meanings with the power to make relation known, that this view posits three powers while related designation only posits two (the designative power, and the power of words to bring about cognition of relation), and is thus more parsimonious. Now from the way in which these arguments are formulated, it would seem that their Prābhākara authors meant them to be taken at face value, as applications of a metaphysical maxim to the effect that, <u>ceteris paribus</u>, the theory which posits fewer entities to account for a phenomenon is preferable. If we assess these arguments in this way, then the related designation theory clearly wins in each case. For in both cases the competing theories are equally capable of accounting for the range of phenomena in question (the cognition of literal sentence meaning), but related designation employs fewer posits than its competitors.

It is not clear, though, that these arguments really ought to be assessed in such a straightforward metaphysical way. In general, parsimony arguments derive their appeal from

the conviction that the phenomena we seek to describe possess an underlying order--and a simpler theory generally depicts the phenomena as possessing a greater degree of orderliness. Hence the maxim that the fewer powers we posit in explaining cognition of sentence meaning, the better. But it is evident that the psychological mechanisms which mediate between perception of utterance and cognition of sentence meaning can bring unwanted complexity to the most metaphysically parsimonious theory. Such complexity can be problematic precisely because among the phenomena to be explained is the fact that language can be learned. We know that our memory and computational capacity have their limits. Thus a theory which buys extreme ontological economy at the price of elaborate and involved mental processing may prove inferior to one which is ontologically profligate but places fewer demands on our mental capabilities. It is possible for a comparative philosopher of language to overreact here, for modern semantics regularly appeals to 'psychological reality' as a way of evaluating semantic theories; and we must bear in mind that the vast majority of non-Buddhist Indian philosophers would agree with Putnam's conclusion that meanings are not in the head. Yet we have already seen instances where the Indian philosophers themselves use parsimony arguments to raise essentially epistemological issues. For instance, the infinite correlates objection, which is couched in terms of the number of designative powers a word must be said to have, really concerns our ability to grasp the designative capacity of a word. Likewise the Prābhākara parsimony argument against the sentence theory gets its force from considerations having to do with our ability to learn. Thus it is not entirely out of place to ask how the competing theories in our present parsimony arguments stand up when the various psychological mechanisms they require are taken into account.

We have already seen that the Prābhākara account of sentence comprehension requires a fair degree of psychological complexity. As we hear a sentence, each word serves to recall its own-meaning together with the fact of its relatedness. This recollection produces a sense of expectancy, or desire to know the remainder, based on our awareness of the incompleteness of the designative function of a word so long as the related correlate is not brought into proximity by another word. Once the utterance is complete, we proceed to assemble the various own-meanings into a relational complex, following the order in which expectancy is generated. This process completed, an act of collective memory yields apprehension of all the own-meanings in mutual relation. Can any of this complexity be avoided if we posit a power-assigning power in words? Indeed, if anything such a posit only makes matters worse. The final stage of collective memory is, as Śālikanātha has pointed out, unavoidable for the designated relation theory,

since word meanings must somehow be gathered together before the mind after they have been individually apprehended. The initial stage of recollection of own-meaning will be replaced by a stage of apprehension of designated word meanings. But between these two stages there must come: (1) the operation of the power-assigning power of words; (2) apprehension of the power thus induced in word meanings to cause cognition of relation; (3) computation of the relations among word meanings in accordance with expectancy; and (4) a check to see if any of the relations arrived at in (3) violate selection restrictions. We can omit (1) if we say instead that word meanings themselves come equipped with the power to bring about cognition of mutual relation. But this claim is unconvincing, since nowhere else does cognition of unrelated entities inevitably lead to cognition of mutual relation. To drop the power-assigning power is to court something akin to the lakṣaṇā thesis, a claim that there is something about language use which signals to us the gap between cognition of unrelated entities and cognition of sentence meaning. And this will mean yet another stage in the processing of a sentence.

The level of psychological complexity required by these two versions of the designated relation theory is not beyond the capacities of humans. But each of these accounts does involve more processing steps than does the related designation theory. And since humans do acquire language with remarkable facility, a theory whose psychological mechanisms are relatively simple would seem to be preferable, given equivalent degrees of ontological economy. The larger lesson contained in all this is that the Prābhākara has gained a more elegant account of the competent speaker's comprehension of sentence meaning by requiring that the language learner grasp more than mere own-meanings when learning the meanings of words. We have already argued that there is no justification for the assumption that the relational element is abstracted out of word meaning when we perform insertion and deletion. Even if our argument on this point is unpersuasive, however, the Prābhākara claim that words designate entities in mutual relation can be justified in terms of the resulting simplification of our account of sentence comprehension. Because we have learned that words designate entities in mutual relation, the occurrence of a word in a sentence will cause us to recall its own-meaning together with the own-meaning's relatedness. The resulting incompleteness of designation generates a sense of expectancy, and this expectancy, together with the expectancies generated by the other words of the sentence, guides our computation of sentence meaning. Considered by itself, this theory has a kind of elegance which is intuitively appealing. But its elegance becomes even more striking when we compare it with the accounts provided by

different formulations of the designated relation theory. Once we see how readily sentence meaning can be gotten from words if words are related designators, it becomes difficult to suppose that there is some element in linguistic cognition distinct from the words that is responsible for our awareness of relation.

The question of the relational element in sentence meaning has not been a major issue in Western philosophy of language. It is, nevertheless, possible to discern three different approaches to the problem of how sentence meaning is constructed out of word meanings. The first is that of the Aristotelian tradition in logic. Its account of word meaning is guided by the metaphor of the relation between a name and its bearer: the meaning of a term is the entity or set of entities it designates. A proposition consists of a subject term, a predicate term, and the copula. Subject and predicate terms name entities, while the copula expresses the characterizing relation that obtains between them. A second approach is that of Frege, who also takes the name-bearer relation as central to the theory of meaning, but abandons the subject-copula-predicate analysis of propositions in favor of an argument-function analysis. Where the Aristotelian tradition would analyze 'Socrates is mortal' into three components, Frege's account yields two: the proper name 'Socrates' and the predicate expression 'is mortal'. Predicate expressions are to be thought of as functions which take proper names as arguments and which have as their values either objects or truth-values. For instance, 'the sister of' is a function which, when given the argument 'Hanuman', names the object 'Abigail'. Likewise 'is the sister of Hanuman' is a predicate expression which, when given the argument 'Abigail', names the truth-value True. Here there is no distinct element in the proposition which expresses relation between subject and predicate. Instead, the sense of a predicate expression is thought of as incomplete or unsaturated, and standing in need of completion by some expression with a sense which is complete, i.e., an expression which names an object.

A third approach is that of Quine, who rejects the use of the name-bearer metaphor in semantics (he refers to it as the 'myth of the museum'), and in effect eliminates the category of proper name. He accomplishes the latter task by assimilating proper names to verbs (which were already treated by Frege as predicate expressions). Thus 'Pegasus does not exist' becomes 'There is nothing which Pegasizes,' and 'Socrates is mortal' becomes 'There is something which Socratizes and which is the only thing which Socratizes, and which is mortal.' Here Frege's argument-function analysis has been modified in such a way that only variables can serve as the arguments of first-level predicate expressions.

Of these three approaches, the first is clearly the least satisfactory. There are, for instance, notorious difficulties involved in extending the subject-copula-predicate analysis to propositions containing verbs other than the copula. It is also noteworthy that of the three accounts, it is only the Aristotelian which in effect ignores the context principle. Frege, on the other hand, was the first philosopher in the Western tradition to give the context principle clear articulation. His argument-function analysis of the proposition, as well as being a significant advance over the traditional subject-predicate analysis, reflects his commitment to that principle. His notion that when we grasp the sense of a predicate expression, the incompleteness of that sense leads quite naturally to our saturating it with the appropriate proper names, reflects his conviction that such expressions can play their allotted semantic role only in conjunction with other expressions. Frege's account has its darker side, though. As we saw earlier, when the 'holes' in a predicate expression are filled with names, the resulting complete expression will sometimes have an object as referent, sometimes a truth-value. This seems rather mysterious. We have also seen that Frege is widely criticized for positing unsaturated objects as the referents of predicate expressions. This posit may well be the result of Frege's having chosen the name-bearer metaphor as the central metaphor in his theory of meaning; but in any event it reflects the tension in his work between the context and composition principles: a firmer commitment to the context principle might have led Frege to the conclusion that predicate expressions do not in fact refer except when their senses are saturated, that is, when they occur in complete sentences. For it seems we are tempted to think that predicate expressions have unsaturated objects as their referents only when we suppose that such expressions must have meaning in isolation.

These difficulties in Frege's views may be avoided if we embrace Quine's analysis of sentences. But this requires us to give up what strikes many as a core intuition in semantics, that the word-meaning relation is in many ways like the name-bearer relation. In this respect the related designation theory represents an important alternative to Quine's way with words. The Prābhākara version of related designation involves a reference theory of meaning, albeit a nonstandard one. Thus the meaning of a word is that which the word designates. But the Prābhākara also, like Quine, extends the argument-function treatment of predicate expressions to all words, and thus avoids the anomaly that complete expressions sometimes name objects and sometimes name truth-values. Moreover, since the theory does not make a radical distinction between proper name and predicate expression (I take it this is what Staal is referring to when he calls the the theory an

extreme form of 'syncategorematicism'), there is less of a temptation to suppose that the latter have incomplete objects as referents. For now even a proper name must stand in a sentential environment in order to succeed in designating its bearer; and so, whatever the core meaning of a predicate expression might be, the name-bearer metaphor no longer leads us to conclude that the referent of such an expression must be just that entity with its relatedness. The referent of any expression is precisely its core meaning in relation to those entities which are the core meanings of those other terms which occur in a sentence in which the expression is used. Thus the related designation theory deserves the attention not only of scholars of Indian philosophy but of those engaged in current work in philosophical semantics as well.

A final point worth noting is this: while it is generally agreed that Indian philosophy of language does not make the sense-reference distinction, there are elements in the Prābhākara account of word meaning that seem to play at least some of the roles allotted to senses by Western philosophers of language. In particular, the notion of a word's own-meaning, which is recalled in the stage of recollection and employed in working out the pattern of mutual relations in sentence meaning, seems rather akin to the concept of a sense as the manner in which the semantic role of a word is apprehended. It also strikes me that the notion of a designative power is rather sense-like in certain respects; but this is clearly a matter which must be pursued in a separate study.[68]

NOTES

1. 'Designated relation' and 'related designation' are easily confused; I wish I knew of a better translation. The designated relation theory (abhihitānvayavāda) is the theory that what is designated by the words of a sentence is subsequently placed in relation. The related designation theory (anvitābhidhānavāda) is the theory that what is designated by words is already related. Thus the two names can be tied to the respective theories if 'designated relation' is thought of as the relation of what is already designated, and 'related designation' is thought of as the designation of what is already related. I have elsewhere used 'the words plus syntax theory' as a name for abhihitānvayavāda, but this I am now convinced is misleading, since the related designation theory also uses syntax in its account of sentence comprehension.

2. Brough (1953).

3. Jha (1918), pp. 61-63.

4. Staal (1966), pp. 116-117.

5. Brough (1953), p. 164.

6. Bhāṭṭacharya (1962), pp. 159-160.

7. Kunjunni Raja (1963), p. 212.

8. Nyāyamañjāri, p. 364.

9. Nyāyamañjarī, p. 367.

10. Nyāyamañjarī, p. 367.

11. Nyāyamañjarī, p. 370.

12. Nyāyamañjarī, pp. 366-367.

13. Nyāyamañjarī, p. 367.

14. Vākyārthamātrkā, p. 379; Tantrarahasya, p. 39.

15. Tantrarahasya, p. 39.

16. Tantrarahasya, p. 39.

17. Vākyārthamātrkā, pp. 380-381; Tantrarahasya, pp.

32-33.

18. Actually, none of the parties would say just this. The Mīmāṃsāka holds that the entity designated by 'cow' is the universal cowness (with or without its relations to other entities). The Naiyāyaika holds that what is designated by 'cow' is a particular inhered in by cowness. We shall ignore this dispute for ease of explication.

19. Ṛjuvimalapañcikā, p. 383.

20. Tantrarahasya, p. 33.

21. Ṛjuvimalapañcikā, p. 383.

22. Vākyārthamātṛkā, p. 400.

23. Ślokavārtika, v. 358.

24. Vākyārthamātṛkā, pp. 392-393; Tantrarahasya, p. 31.

25. Vākyārthamātṛkā, p. 394.

26. Mānameyodaya, p. 98.

27. Vākyārthamātṛkā, p. 392; Tantrarahasya, pp. 30-1.

28. Tarkasaṁgrahadīpikā on Tarkasaṁgraha p. 66.

29. Tantrarahasya, p. 30.

30. Tantrarahasya, pp. 30-31.

31. Vākyārthamātṛkā, p. 385.

32. Tantrarahasya, p. 33.

33. Tantrarahasya, p. 33.

34. Vākyārthamātṛkā, p. 386.

35. Vākyārthamātṛkā, p. 386.

36. Mānameyodaya, p. 96.

37. Vākyārthamātṛkā, pp. 397-398; Tantrarahasya, pp. 31-32.

38. Vākyārthamātṛkā, pp. 398-399.

39. Tantrarahasya, p. 32.

40. 'Was Śāntarakṣita a "Positivist"?' forthcoming in Buddhist Logic and Epistemology, ed. B.K. Matilal, D. Reidel, Dordrecht.

41. Vākyārthamātṛkā, p. 400.

42. Vākyārthamātṛkā, p. 400.

43. Vākyārthamātṛkā, pp. 381-383.

44. Ṛjuvimalapañcikā, p. 384.

45. Vākyārthamātṛkā, p. 406; Tantrarahasya, p. 38.

46. Vākyārthamātṛkā, p. 405, emphasis added.

47. Tantrarahasya, p. 39.

48. Tantrarahasya, p. 39.

49. Vākyārthamātṛkā, p. 402.

50. Tantrarahasya, p. 38.

51. Vākyārthamātṛkā, p. 408.

52. Tantrarahasya, pp. 34, 38.

53. Matilal (1966) discusses the dispute between the Naiyāyikas and Grammarians over the structure of a sentence (see especially pp. 388-392). If, as seems likely, this dispute cannot be resolved, then the rules of sentence structure will turn out to be more complex than our Prābhākara authors suppose.

54. Tantrarahasya, p. 39.

55. Tantrarahasya, p. 39.

56. Ṛjuvimalapañcikā, p. 384.

57. Vākyārthamātṛkā, p. 402.

58. Ṛjuvimalapañcikā, p. 383.

59. Vākyārthamātṛkā, p. 381.

60. Vākyārthamātṛkā, p. 394.

61. Vākyārthamātṛkā, p. 394.

62. Here I am assuming that the inherence which relates colorness and a particular green, and the inherence which relates colorness and a particular brown, are distinct entities. This assumption would not be granted by a Naiyāyika, but it is adopted here solely for purposes of illustration.

63. Nyāyamañjarī, pp. 368-369.

64. As Mohanty (1966) describes the dispute over validity, this is really the position of the Bhāṭṭas, not that of the Prābhākaras (see especially pp. 6-23). The latter, according to Mohanty, maintain that strictly speaking there is no real error; all presentative apprehensions are valid insofar as they succeed in presenting their objects to consciousness. Jayanta does not here explicitly ascribe the extrinsic invalidity view to the Prābhākara, but I see no way of making sense of his remarks other than to suppose him to attribute this view to them.

65. Nyāyamañjarī, p. 369.

66. This is how Rāmānujācārya defines fitness at Tantrarahasya, p. 35.

67. Nyāyakośa, p. 674.

68. Research for this paper was begun during an NEH Summer Seminary for College Teachers on the topic of formal semantics and natural language, conducted by Professor Richard Grandy at Rice University in the summer of 1982. This paper was completed with the support of an NEH Summer Research Stipend for the summer of 1983. The support of the National Endowment for the Humanities is greatly appreciated. The following individuals contributed helpful comments and criticism during various stages of this project: Richard Grandy, J.N. Mohanty, Bimal K. Matilal, and Kenton F. Machina.

REFERENCES

Annambhaṭṭa: 1976, Tarkasaṁgrahadīpikā on Tarkasaṁagraha, edited and translated by Gopinath Bhattacharya, Progressive Publishers, Calcutta.

Bhattacharya, Bishnupada: 1962, A Study in Language and Meaning, Progressive Publishers, Calcutta.

Brough, John: 1953, 'Some Indian Theories of Meaning', Transactions of the Philological Society 1953, 161-176.

Jayanta Bhatta: 1936, Nyāyamañjarī, (ed.) Sūrya Nārāyana Śukla, Kashi Sanskrit Series 106, Benares.

Jha, M. Ganganatha: 1918, The Prābhākara School of Pūrva Mīmāmsā, Benares Hindu University, Benares.

Jhalakīkar, M. Bhīmācārya: 1978, Nyāyakośa, Bombay Sanskrit and Prakrit Series XLIX, Bhandarkar Oriental Research Institute, Poona.

Kunjunni Raja, K.: 1963, Indian Theories of Meaning, Vasanta Press, Madras.

Matilal, Bimal K.: 1966, 'Indian Theorists on the Nature of the Sentence (vākya)', Foundations of Language 2, 377-393.

Mohanty, J.N.: 1966, Gaṅgesa's Theory of Truth, Center of Advanced Study in Philosophy, Santiniketan.

Nārāyana: 1975, Mānameyodaya, C. Kunhan Raja and S.S. Suryanarayana Sastri (eds. and trs.), Aydar Library Series 105, Vasanta Press, Adyar.

Rāmānujācārya: 1923, Tantrarahasya, R. Shamashastri (ed.), Gaekwad Oriental Series XXIV, Central Library, Baroda.

Śālikanātha: 1934, Rjuvimālapañcikā, with Prābhākara's Brhatī, S.K. Sastri (ed.), Madras University Sanskrit Series 3, Madras University, Madras.

Śālikanātha: 1954, Vākyārthamātrkā, in Prakarana Pañcikā of Śālikanātha, Subrahmanya Sastri (ed.), Banaras Hindu University Darśana Series 4, Benares Hindu University, Benares.

Staal, J.F.: 1976, 'Sanskrit Philosophy of Language', in Herman Parret (ed.), History of Linguistic Thought and Contemporary Linguistics, Walter de Gruyter, Berlin, pp. 102-136.

A. Chakrabarti

PLATO'S INDIAN BARBERS

> this is the old Platonic riddle of nonbeing. Nonbeing must in some sense be, otherwise what is it that there is not? This tangled doctrine might be nick-named Plato's beard; historically it has proved tough, frequently dulling the edge of Occam's razor.
>
> Quine
> (From a Logical Point of View (1963), p. 2)

1.

From the numerous contexts in which the status of empty terms has been discussed and debated upon in different periods and in works of different schools of Indian Philosophy, I have chosen only two for detailed discussion, the Nyāya and the Grammarians'.

The Nyāya school of direct realism that has the richest tradition of logical and epistemological analysis running in an unbroken chain from Gautama (1st century AD) to Gadādhara (16th century AD) propounds that the class of things which are real, the class of things which are correctly cognisable and the class of things which are genuinely designated by words have strictly the same extension.

"We do not recognize a single term without a denotation."[1]

Uddyotakara (6th century AD) who made the above announcement follows the "deflationist"[2] policy of paraphrasing away all terminus fictus in favour of simpler parts referring to real objects in the typical Ockhamian fashion[3] when he discusses the standard example "The rabbit horn does not exist."[4] I have translated and examined his discussion on this point elsewhere.[5] Deriving the basic insights from Uddyotakara, Udayana--the 11th century consolidator of the Nyāya-Vaiśeṣika logico-metaphysical doctrines--engages in a protracted polemic with the Buddhists concerning empty subject terms of negative existentials, part of which I reproduce in the next section.

B. K. Matilal and J. L. Shaw (eds.), Analytical Philosophy in Comparative Perspective, 299–326.

2. UDAYANA VERSUS THE BUDDHISTS

To exist is, for the Buddhists, to be causally capable. Since they elaborately argue against the dispositional notion of potential capacity, for them a thing exists as long as it actually produces something. Production is taken strictly as an instantaneous act, finally boiling down to "replacement by a next bit of reality". The causal predecessor gives birth to its own real-making offspring by dying as fast as possible. Hence, the hallowed doctrine of momentariness. The resemblance of their activity criterion of existence with Frege's notion of _Wirklichkeit_ is only apparent. Something unchangeable would not be called _Wirklich_; but Frege would not require that the actual be so change-ful as not to endure long enough to be the sub-strate of a change. The Buddhists have to make strong statements like the following: "The non-momentary is nonexistent"--and, thus, get involved in the problem of referring to the non-momentary for denying existence and productivity of it.

For Frege, the notion of actuality as a property of _some_ individuals did not create any parallel problem because he had the wider (and higher level) concept of being there or being given. So he could easily say, without any fear of inconsistency, "There _is_ an _x_ such that _x_ does not act upon anything and is hence not actually real". But to the Buddhist no such wider and presupposable notion of being was _prima facie_ available; therefore, they would be hard put to point at a non-productive, hence, non-real object. It is in this setting that Udayana picks up his logical quarrel with the Buddhists' negative existentials like "The causally barren entity does not exist".

To explain this, a little background information is required. In the typically Indian logical model of inference[6] we have to back up the generalized statement of "pervasion relaion" between the mark (_hetu_) and the inferable (_sādhya_) by positive and negative instances. A statement of pervasion is corroborated by an instance where both the mark and the inferable co-occurs, and is strengthened by the absence of the mark in every place where the inferable is surely known to be absent. This latter sort of case, i.e. cases of sure absence of the inferable are called "disagreeing examples" (_vipakṣa_) and showing the non-occurrence of the mark in a disagreeing example is considered to be of great importance for the establishment of the _mark_ as a conclusive one. Thus, unless one is sure that any arbitrary instance of a non-communist would be a non-Marxist, one cannot take being a Marxist as a good reason or "sure mark" for inferring that someone is a communist. In other words, we must find the truth of a universal instantiation of the contra-positive of the generalisation, as well as

an instantiation of the generalisation itself. But in some inferences which concern everything (i.e. everything becomes the paksa or subject of the conclusion of the inference) and purport to prove the presence of a universal inferable property from--obviously--a universal reason or mark, it is best to keep silent about whether the mark is present or absent in the disagreeing examples because such examples would be nonexistent.[7] The Buddhists flout this warning when they claim to have supported the generalisation "Whatever exists is productive" by showing the absence of the mark--productivity--in the disagreeing example of a nonexistent object, e.g. a barren woman's son. Quite predictably, Udayana clinches the argument at this point.[8]

Among the enumerated defects of inference, Indian logic counts the fallacy of fictitious subject, which is classically illustrated by:

> The sky-rose is fragrant,
> because it is a rose, like the rose in the garden.

Since, according to the Buddhist, the "non-momentary" is a vacuous term, "...lacks activity" is an empty predicate and "the rabbit-horn", an avowed non-entity, all the three essential elements of inference--the subject, the mark and the example are "unestablished" in the alleged inference

> The non-momentary is nonexistent
> Because of its lack of activity, like the rabbit-horn.

Udayana remarks:

> There cannot be any means of knowledge to establish a non-entity. If it could be established by some means of knowledge it would cease to be a non-entity.
>
> Ātma Tattva Viveka, p. 59.

In reply the Buddhists echo the classical retort of the Sophist:

> The sentence by which you negate the possibility of the nonexistent being a bearer of properties itself ascribes the property of "lack of all properties" to the nonexistent.[9]
>
> Kṣaṇabhaṅgasiddhi Vyatirekātmikā

This reply of Ratnakīrtī is again based on Jñānaśri's remark:

> Can there be any provable (epistemically justifiable) affirmative or negative attribution to a non-entity? If, in some cases, such predication is admitted as valid, how do you account for only those cases being free of logical defect? If no such predication is allowed then (by asserting the self-referring proposition that non-entities are undiscussable) you contradict your own contention.[10]

In the face of this charge Udayana seems to concede, at least provisionally, that in making the prohibitive statement about non-entities, he was incurring a kind of self-refutation. Later on, in the same treatise, we come across a highly perceptive general account of the types of inconsistency, self-refutation and awkwardness that can be manifested in a sentence. There he makes fairly fine distinction between

(a) Self-defeating statement, e.g. "My mother is childless".
(b) Self-defeating act, e.g. "I am mute" when uttered.
(c) Self-defeating cognition, e.g. "I do not cognize this" when expressing a cognition.

Here, though the self-refutation concerned is said to be of type (a), it seems rather loosely classified as so.

Self-referential paradoxes were well-known to ancient Indian logicians. Bhartṛhari[11] discusses a series of such paradoxes, e.g. "All that I speak is false" or "No contentions can prove anything" or "This object is not talkable about". Like the standard way out, Bhartṛhari's general solution, too, is that when we are making such a prohibition we suppose our own assertion to be of a different level and the negation "to be applicable only to statements other than this one".[12]

But, Udayana, instead of resorting to any such self-exempting reply, courts the more challenging position that there is indeed such a self-refutation involved in his own statement that the non-existent rabbit-horn is not the subject of any intelligible discourse. We can react in three alternative ways to such a stalemate.

(1) We might think that the fact that the denial of discussability of the nonexistent results in contradiction constitutes an indirect proof of the admissibility of the nonexistent as a topic of talk.
(2) We might abandon all hope of discovering the logic of our discourse about non-entities

and desist from giving any decision on the matter.

(3) We might reject both affirmative and negative predication which claim to be irreducibly about an unestablished non-real entity because both sorts of statements would be equally unfounded.

Udayana embraces (3). But doesn't (3) force him to total silence over this subject rather than his prohibitive diatribe against it?--asks the opponent. Yes, it does. Faced with the question: "Is the rabbit-horn sharp?" one might either keep mum or one can rave about resorting to circuitous ways of answering yes or no. And who would be the wiser between the two? Udayana, a master of style, leaves this question for the reader.

He then enters into the heart of this terrible tangle where silence is prescribed as the panacea but only through the admittedly self-baffling talk.

Reflecting on the following proposition:

> The non-existent cannot be made the subject of any predication... (N_1)

the Buddhists think that at least such negative assertions as N_1 can be made regarding the non-existent. Hence,

> The non-existent is a possible subject of negative predication... (N_2)

Udayana could have easily broken through the logical knot by protesting that N_1 is not a statement about <u>the non-existent</u> but is only about <u>statements</u> which do not have any existent entitiy as their subject, and it says about such statements that they are neither true nor false, but do not deserve logical appraisal. Instead, he does admit that N_1 ascribes untalkability-about (<u>avācyatva</u>) to the non-existent[13] and is only at pains to point out that even if N_1 displays its own futility, that does not build any justification for accepting

> The non-existent is unproductive... (N_3)

as either true or false. The self-baffling nature of N_1 does not allay the worry that once we admit even the so-called "impredicative"[14] assertions as correctly <u>about</u> what is not there, we shall lose all logical protection against "wandering thoughts" and arbitrary intuitions, because there is no objective method of verifying those assertions, since their topic remains admittedly inaccessible to all trustworthy cognition.

Here the Buddhists appeal to such ordinary language statements as "The sky-lotus is not real", "The childless woman's son is an impossibility". Belief in their truth seems to be deeply rooted in our linguistic practice.

In the course of the entire debate, the Buddhist is never represented as appealing to such affirmative predications such as

> The golden mountain is a mountain.

Rather unlike the vulgar Meinongian, he would refuse to allow the following analytic looking sentence to be true in any sense:

> The causally potent sky-lotus is causally potent.

All these properties--most probably all positive properties--which are exemplified in real entities are held to be exclusive attributes of existent objects (vastu dharma) whereas negations of them were mostly taken in their widest sense. Thus, "non-red" to them would not stand for the property of being of some other colour (as it would, perhaps, to Udayana) but would denote a concept under which falls anything provided it is false to say of it that it is red. Ratnakīrti would have fully agreed with Routley in holding that "... non-entities have definite properties";[15] and also would admit that it is true of Unicorns that "...unicorns are not the sorts of items that are prime" but would not say, like him, that they have horns or that they are animals. In this respect, they are much more hard-headed than Meinongians. Their insistence that it must be granted as true that the rabbit-horn is not sharp is based upon the same reasons that prompted Russell to allow that there is a sense in which "The present King of France is not bald" is true. Taking negation in the sense of exclusion-negation rather than choice-negation, the Buddhists would assign truth to any sentence which denies a certain property of a non-existent object. Both the Buddhists and Udayana would agree that a sentence like, "Wittgenstein's wife was tall" would be unserious and truth-valueless. But while its negation would be taken to be just as inane by Udayana, the Buddhists would take "Wittgenstein's wife was not tall" (W) not to mean that she was short or of medium height but to mean the same as "It is not true that Wittgenstein's wife was tall" (W') and call it true in that sense. That would have been all right, as Matilal wants to show,[16] by the three-valued truth table for exclusion negation and his additional speech-act operator "It is true that" which yields falsehood if the embedded sentence is either false or indeterminate or senseless. But the Buddhists make another

move at the same time. They jump from the truth of W, forgetting, as it were, the secondary occurrence of the empty definite description "Wittgenstein's wife" in W', to the ontological observation that the non-existent object fitting that description really possesses the property of not being tall. But this property of non-tallness is never identical with that of being short or of normal height. Ratnakīrti remarks:

> Three sorts of properties are found. Some sit always on real objects, e.g. the colour blue. Some, as a rule, go with unreal objects, e.g. total lack of specifiability. Some again are found in real and unreal objects, e.g. mere non-apprehension [both an existent and a non-existent black cat may remain unseen in a dark room].
>
> McDermott, p. 19.

It is agreed that some negative properties, in fact most of them, belong both to real and unreal objects, but there are some like lack of activity, lack of momentariness and lack of positive description which are true only of non-existents. To quote Ratnakīrti again

> We may say in this context that just as we find usage of words in the manner of property and propertied about real entities, e.g. cowness in a cow, whiteness in a cloth, running in a horse, the same property-propertied mode of speaking is also found about unreal objects, e.g. lack of sharpness in a rabbit horn, absence of speakerhood in a childless woman's son, the scentlessness of a sky-lotus etc.

The general Buddhist position is, therefore, very un-Meinongian insofar as it does not deduce the predicate-worthiness of non-entities from what Routley calls the characterization postulate $\vdash (\iota \underline{x})\phi\underline{x}$ is ϕ.

There is an old grammatical distinction used in Indian logic between external and internal negation which the Buddhists seem to rely on. It was initially formulated (somewhat like the distinction so often appealed to by Meinong and his followers, between sentence negation and predicate negation) as the contrast between "not" governing the main verb and "not" governing the noun. Consider

A does not speak .... (S)
A is a non-speaker .. (S')

The grammarians thought that in S type of negation the force of the assertion is merely that of disavowal, with no commitment to any corresponding affirmation of the complementary predicate. In S', implicit ascription of the complement predicate to A dominates. The first is called, literally, repelled suggestion and the second, positing of the contradictory.[17] Udayana takes negation in the second committal way and construes "A is a non-speaker" as "A _does_ something other than speaking" (_A.T.V._, p. 68).

Udayana objects to a joint adoption of both the expedients (a) of taking the negation as an external, non-committal sentence negation and (b) of taking _such_ a negative sentence as evidence that we do talk _about_ and apply predicates to non-entities. To translate the bit of dialogue which follows:

> If you insist that a non-entity can be the subject of a denial, we would ask: Why can't it be the subject of an affirmation too? The absence of valid means of knowledge remains the same in both the cases.
> _Opponent_: That the son of a childless woman does not speak can be established by the mark that he is not conscious [note the conniving use of mark-to-inferable pattern of argument even concerning an acknowledged non-subject] but there is no way of proving that the son of a childless woman does speak [thus we prefer the negative form of statement].
> _Proponent_: No. Even to prove that the son of a childless woman _does_ speak you can assign the reason that he is a son. You cannot say that a childless woman's son is not a son, because if you do so, you will contradict yourself.
>
> _A.T.V._, p. 65.

This last point brings us to the most interesting use that Udayana--surprisingly--makes of the Meinongian characterization principle. He surely does not hold the position himself that the son of a childless mother really belongs to the class of sons. (Did even Meinong seriously believe so?)

Although the dispute here has a tendency to lose itself in sophistry, certain interesting and profound differences of the basic _Nyāya_ and Buddhist semantic outlook have come to the surface in this dialogue.

Udayana, too, would have no doubt assented to the _de dicto_ claim that it is not the case that speakerhood belongs to the childless woman's son. But the moment it is taken _de re_ by the Buddhists as a case of a presence of speechlessness _in_ an object of disingenuous reference, he would ask the

following very pertinent question

> Do you say this, honestly, because you have regularly apprehended a non-entity as devoid of speakerhood, or because you have simply failed to apprehend speakerhood anywhere outside real objects?[18]
>
> A.I.V., p. 67.

It is only with a mock sympathy with common usage of empty terms in language that the Buddhists say that, after all, we do talk intelligibly about non-entities. They first use the Russellian device of negation governed utterances "about" non-things like tortoise-wool, but then use that phoney topichood of unreal objects to construct a superficial Meinongian ontology of given fictions (vikalpas). The tendency is to blur the distinction between empirically real, stable objects like hills and houses, pots and people on the one hand and figments of imaginaion on the other. Hence, the identification of property bearers with property repellers, which suits them quite well since, in their polemic against the reality of universals, they construe all subsumption under F as exclusion from non-F. Their ontology is, in a way, like that of Meinong seen in a photographic negative. While Meinong objectifies non-existents as independent of our awareness, the Buddhists relegate the conventionally objective substances into mere linguistic and cognitional fictions. The incoherence in the Buddhists position that Udayana labours to pinpoint becomes clear when you notice how it tries to lend support to a proto-Meinongian ontology by justifying negative assertions with vacuous subject-terms with a Russellian trick over the scope of negation. A spurious object like the rabbit-horn is not fit to hold even a real absence. If it is said to be non-sharp simply because it is not there to be sharp, that is surely not its being non-sharp. As we have seen earlier, the Nyāya takes absences so seriously that without definite checkable information about its absentee's residence, it would never accept a particular absence.

In course of debate Udayana draws his opponent into a circular reasoning. When the negative attribution of speechlessness is construed by Udayana as a positive ascription of agency to something other than speech, the Buddhists protest that this cannot be done because the non-entity is devoid of all agency. How does he establish that the non-entity cannot act? From the simple premise that it is not existent. But what is the reason for failing to exist? The fact that it cannot act to be sure.[19]

Such a circularity, of course, might not be too frightening to the Buddhists because they expressly take the statement

"A non-entity is inactive" as analytic and definitional. But then, what Udayana decries is the pretension of proving, as if through inductive argument based on instances, that whatever does not exist is found to be causally fallow.

Udayana then issues his warning:

> The limits of valid cognition are the limits of [logically appraisable] linguistic behaviour, beyond which there is no rule [as to truth conditions].
>
> A.T.V., p. 69.

The risk of rulelessness (aniyama) looms as the main resistance that Udayana feels against non-entities entering the domain of linguistic or cognitive reference. We do not genuinely understand a statement about a non-entity because we have no definite means to say when such a statement would be true and when false. Empty predictes are inadmissable because their application is not directly learnable (except in terms of other non-empty predicates) and also because there is no objective way of distinguishing the ranges of two such different predicates (like "... is a bandersnatch" and "... is a flibbertigibbet") without reducing them to their instantiated constitutents.

In a strikingly Stawsonian tone Udayana remarks:

> About an ever-unapprehended [fabricated] Devadatta, the question "Is he fair or dark?" does not arise except in a spirit of wantonness. And if, without caring to understand what this is all about, someone answers "Fair" why shouldn't another give the answer "Dark"?
>
> A.T.V., p. 69.

Since the double defects of (a) absence of evidence and (b) inconsistency (insofar as the alleged answer about his complexion implicates an already disowned knowledge of the existence of the person) equally infect both the contrary answers, neither of them can be preferred to the other.

Udayana now anticipates another epistemological defence of the Buddhist position in favour of non-actual intentional objects. It is not correct to maintain that we talk only about what we have knowledge of. Don't we talk about the content of our illusions, the horrifying monster I hallucinated the other night or the tortoise hair that a child seemed to see from a distance (mistaking the grooves on the shell for hair)? When cognitions of such figments arise, they come to us with a clear differentiation of their objects. We do

distinguish between the illusory horned rabbit from, say, the illusory shadow-trick ghost. Sometimes we can even re-identify them within the short duration of an illusory experience. In a dream event we make repeated back-references to the same person, the same room--and guard ourselves carefully from confusing one item with the other while relating our just-dreamt dreams to someone else. In all these contexts we keep consciously talking about things not known to be existing or happening but perceived falsely, and people do understand our talk as about such things with full knowledge of their non-existence.

Udayana rejects the above consideration using the following argument. Let us suppose we sometimes do have some sort of apprehension referring to a rabbit's horn. What should be the precise analysis of such a cognition? Shall we take it as a case of otherwise-apprehension or a case of unreal-apprehension?

The former alternative is not palatable to the idealist Buddhist, in search for an example of a single non-real object of mind; for, upon such a construal, there must be something (real and elsewhere perceived) which is superimposed and something on which it is superimposed. The place of superimposition must be there (accessible to demonstrative reference, e.g. of the subject expression of the false proposition: "That is a horn") and what is superimposed must also be existing somewhere else and, thus, you are won over by the Naiyāyikas. (A.T.V., p. 70).

Obviously then, the Buddhists have to resort to the proto-mādhyamika"[20] theory of error, according to which an unreal object appears as real in error. Here Udayana argues rather cursorily against a somewhat weakly represented position that the cognition of rabbit-horn is best described as awareness of a non-existent individual with the false belief that it is existent. What sort of cognition would it be--he asks--perceptual, inferential or due to understanding of uttered words?

It cannot be perceptual because the sense organs generate perception only when stimulated by some external objects. In perceptual awareness, the mind is merely receptive; it is the object which dominates. A completely imaginary object cannot lie out there to activate our sensory faculties. If a rabbit-horn could come into contact with my senses, it would no longer remain an unreal object. When, let us imagine, a child "sees" the horns of a rabbit, what actually stimulates his visual sense may be the straightened, upturned long ears of the rabbit, which may arouse memory traces of an earlier perception of horns in other animals. The mark of pastness lapses from the object of memory and the remembered property of hornness gets imputed upon the immediately perceived objects. The objects (the ears) are there, but they are seen

as horns. This seeing as can be easily, and, in fact, better, explained without positing a direct intentional object like the seen but non-existent rabbit-horn.

Even fallacious reasoning and linguistic cognition of false sentences can give rise only to otherwise-apprehension, insofar as such cognitions, unlike some sensory awareness, come always with a ready judgmental structure. If we say that we do sometimes chance to misunderstand people and grasp an unintended meaning from a word, can't we, in this fashion, after hearing the words "a rabbit's horn" come to cognise a single entity which is neither the meaning of "rabbit" nor of "horn" nor of any referring expression whatsoever, but is a distinct non-existent item?

In that case, Udayana perceptively remarks, there is nothing which can stop me from understanding tortoise-wool or some other imaginary object by the same expression, viz. "a rabbit's horn". Once we have given up the conventional connection between words of a language and extra-linguistic entities of the world, and deviated from the systematic way of constructing meaning of compound expressions on its basis, we cannot have any rule which can make us pick out exactly one precise intended non-entity and not another.

This reveals a further extensionalist trait of Udayana's thought. The want of definite criteria of identification and differentiation is the main obstacle to the admission of single designata of empty terms. Semantic rules assign real individuals to names and non-null sets of objects to predicates. There is no rule which can fix the referential links between distinct, vacuous singular terms and discriminable non-existents simply because one end of the link is always missing.

Yet, we must do justice to the intuitively evident nonsubstitutability of empty descriptions. A funny couplet is quoted by the Sanskrit Grammarians:

> Here goes the barren woman's son, his hair adorned with a sky-grown flower,
> Brandishing a bow made of rabbit-horn and clad in a tortoise-fur jacket.

However absurd this may sound, there is a certain suitability of the particular empty terms for their particular positions in their fantastic description. That is best explained by the presence of non-empty substance-words at the end of every vacuous descriptive phrase. We know that in real life sons can walk, flowers adorn, horns may be used to make bows and fur for jackets. The quasi-understanding that we now have of it will be blocked if we reshuffle the empty terms as follows:

> Here goes the tortoise-wool, his hair adorned
> with a barren woman's son...

The explanation of the fact that we do not confuse rabbit-horns with sky-flowers despite recognising the nullity of both is not to be sought in our coming to grasp distinct unreals but in the meanings of the non-vacuous consitituent words "rabbit", "horn", "sky" and "flower". Thus, even a verbal cognition of the meaning of "rabbit-horn" does not require a non-existent item as its simple object.

Let us suppose that there is some determining factor which regulates our understanding of these vacuous expressions, something that compels us, as a rule, to grasp a rabbit's horn on hearing the words "a rabbit's horn" and to understand a mermaid by the word "mermaid". Arranging the argument in his favourite form of a destructive dilemma, Udayana proceeds as follows.

Such a hypothetical regulating factor can be either a learnable semantic convention (a rule of the form "By 'w' is to be understood w") or the very intrinsic nature of the word by virtue of which it is bound to arouse in us the appropriate idea or make us grasp its proper meaning.

On the former supposition, the semantic rule can be given either for the entire compound word "rabbit-horn" or separately for the component words. It cannot be given for the whole word, because we would then have to be in the possible situation where the instructor points to a rabbit-horn and says "This object and its like is to be understood by the word 'rabbit-horn'". In the basic Nyāya picture of learning the meaning of a simple word, we have to perceive or inferentially cognize the denotation or a sample of the denotation of the word alongside the word itself, at least once in the beginning. As Kripke and others maintain for proper names and natural-kind terms, the Nyāya believes that every simple word gains its meaning by sustaining a link with the original (humanly or superhumanly) established connection between the word and what it refers to . If it is theoretically impossible to be in the presence of the referent, the word, by itself, can have no independent meaning.

To maintain that we do cognise the rabbit-horn in some way, viz. as the designatum of the word "rabbit-horn" would be of no use. A lexical entry of the following form:

> By "flibbertygibbet" is to be understood that
> which is denoted by "flibbertygibbet"

would be quite unilluminating. If that is the only way to give the sense of a word then it amounts to the word having no sense at all which is not true in the case of the word "rabbit-horn".

If, on the other hand, the semantic rule is given separately for the component words, the words "rabbit" and "horn" need not give up their own conventional meanings. The process of understanding these unsatisfied descriptions becomes as simple as this: First, grasp the meanings of the constitutent words separately and then think of them as related according to the pattern suggested by the compound with or without the recognition that they are not actually so related in the world. But this falls squarely into the model of otherwise-apprehension, which renders a separate designatum for vacuous singular expressions redundant (A.T.V., p. 72).[21]

Let us now consider the other horn of the dilemma. Does the word connect itself to its own empty reference by its own nature (svabhāva)? If a word like "khapuspa" could, by its essential (phonetic?) power evoke the required meaning, then like the person who knows a language (e.g. Sanskrit), the person who does not know it would also have had the same idea of the fictional sky-flower on simply hearing the word (ibid., p. 72).

Udayana then critically considers an alternative intentionalist account of imaginary objects. Although a unitary unreal item can neither be perceived by the senses nor be designated by an empty term, can't we grasp the concept of a rabbit-horn as that which another person has in mind when he utters the word "rabbit-horn" with the obvious intention of conveying some message?

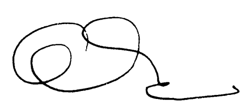
Simply by recognising the speaker's honest intention to refer by the use of the following expression

"a flostrophobous groose"

can we understand anything at all, if no part of that gibberish conveys any conventional non-empty meaning to us? It would be no help for our imagination if that indefinite description is accompanied by the following illustration



A groose



How a groose looks when it is flostrophobous

unless we associate every part of those jottings, systematically, with the looks of some familiar object of our experience. This is the basic principle of an empiricist theory of meaning. If semantic reference is completely lacking we cannot build our understanding in vacuum simply on the basis of speaker's reference. Speaker's meaning (tātparya) can

touch-up or modify word meaning (_śakti_) but cannot replace it. At some point the story teller's world of words must touch the world of experience and knowledge for the story to make sense.

Finally, Udayana anticipates a more subtle form of the above intention-mentioning account of non-existent reference:

> Well, it is easy to learn the meaning rule for such a word in the following manner: There is _some_ designatory intention of the utterer of the word "rabbit-horn" and the object of that intention is what the word speaks of.[22]
>
> (_A.T.V._, p. 75).

Udayana rejects this by appealing to what he takes to be the standard way of grasping the meaning of a new word. When we first hear the sentence "Rope the cow" we do not rest satisfied by telling ourselves that those sounds must have some meaning that their utterer intends them to have, without bothering to look around for an object or an activity with which the distinguishable parts of that string of noise may correspond. The language instruction which always keeps us afloat at the level of "What the speaker intends to mean by those sounds" gets us nowhere. We never _know_ a language by merely referring to some unidentified indefinite objects of a wish to refer which we believe accompanies others' use of words.

At this juncture another doctrine of understanding is provisionally attributed to the Buddhists. According to it, the fictional objects are just creations of our own individual emotions and other coloring effects of memory and desire. The child wishes to ride a horse so strongly that it finds a live horse in a stick with a curved handle, not looking even remotely like a horse's head. A man madly afraid of murder sees a dagger in the air. Our desires and fears cook up fantasies and dreams which provide referents for our vacuous terms.

But then your chain of desires will create an entity distinct from mine, and the two creatures would inhabit completely different worlds. Those worlds being private and insulated ones, you will know nothing definite about my object of fantastic reference, and I will have, at best, a guess as to what your objects of reference are.

> How can two men, completely ignorant of each other's contents of information, communicate about the nature of those contents?
>
> _A.T.V._, p. 76

Our talk about such fictional items would then never be meant for others--and Udayana hints that such a private language would not be a language.

To make someone else understand what I mean by a word, I must be able either

(a) To publicly handle the object(s) denoted, or

(b) To demonstrate the object as that yonder one, or

(c) To point out its perceptible similarity with another tangible familiar object, or

(d) To describe it truly by using words with known meanings, thereby stating an objectively observable situation.

Since we can do none of these with a creature of private fancy, it does not fare at all well as a proposed unitary referent of a term like "rabbit-horn". What we deny the existence of cannot, and need not be such an object of private desire.[23]

Unlike the classical Indian Grammarians who built up thought objects as primary meanings of words, the Nyāya would have us analyse all verbal expressions to simpler components until we reach elements that directly designate real items available to experience in the wide Nyāya conception of the term.

The Nyāya does recognise, somewhat like Frege, that strictly empty singular terms may be used in a sentence which can be mock-understood. A sentence like "The white toy elephant ran to the teddy bear and said..." gives rise not to verbal knowledge but to what they call quasi-understanding (āhāryajñāna). Both our metaphorical and fictional understanding of elements lacking in straightforward semantic congruence have been covered by this notion of quasi-understanding which has been sometimes defined as "an awareness or inner perception which is due to desire to take as true during a full recognition of contrariety to fact".

3. BHARTṚHARI

> Merely on the basis of words heard men are found to apply the category of a thing, even to totally unreal objects like the so-called cinder cycle.[24]
>
> Vākya Padīya, 1/129

Grammarian philosophers have traditionally believed in intentional entities which are created by competent language users' genuine wish to refer by words, even when the words are known to be directly inapplicable to anything in the external world. We have chosen Bhartṛhari (3rd century AD) as their representative. There have been other later thinkers in this tradition who have clearly stated that the primary meaning of an individual word is not an external object but a vikalpa. It is not at all clear what exactly they mean by this key term which is found in the Yogasūtra of Patañjali, 1.9, where it is defined as "a mental construction which is devoid of a corresponding object and is conjured by cognitions arising from words". Major Buddhist philosophers like Dinnāga had borrowed this term from Bhartṛhari. The word has been variously translated as "imaginary objects", "conceptual constructs", "logical fictions" or "thought posits" with different shades of theoretical bias.

In Buddhist philsophy these thought posits are contrasted with directly presented, immediate, atomic objects of sensory acquaintance--the genuine uncoloured bits of fleeting reality. In the Vedantic interpretation of Bhartṛhari's philsophy, vikalpas are the stuff of which the false phenomenal world of multiplicity is built up on the true foundation of the One Absolute. It is by the vikalpas that we divide this undifferentiated reality, the Substance (this is the metaphysical notion of substance other than the day-to-day practical notion (see Matilal (1971), p. 110) to provide for the referents of our individual words. Bhartṛhari is famous for his doctrine of the unreality of the words. To him words are artificially chopped out of the indivisible body of the sentences which alone have meaning in the operation level. They are constructs of grammatical analysis. Word meanings, too, thus, belong to the domain of description and dissection of language by the grammatical categories and have no intrinsic actuality. The one ultimate reference of language as such--the Reality to which all our sentences are predicates (a notion F.H. Bradley would have welcomed)--is not by itself describable in a structured language. Whenever we talk about it with words we slice it into vikalpas. We look at it through the conceptual categories, as if through narrow cylindrical pipes (V.P. III/2/verse 8) which impose pragmatically useful boundaries in all pervading ultimate referent of language. It is interesting to note that the same analogy of looking at the same object through different telescopes occurs in Frege when he compares the senses of words with the real images cast by the moon upon lenses of different telescopes (Frege (1980), p. 61). The two contexts, of course, were quite different and Frege was no Monist in his ontology.

Helārāja gives a more or less clear description of how words mean according to Bhartṛhari.

> When a word is uttered there arises a piece of cognition endowed with a particular form. It is that form which is the meaning of the word--not that act of cognition--which is completely subjective [compare Frege's distinction between evoked _ideas_ and expressed _senses_; the word "form" need not be taken too literally--it can be taken as some sort of blueprint for an object]. Whether that form is ["is" may mean "fits" or "matches" if we stick to our blueprint interpretation] an external object or not does not make any difference to its being the meaning of the word.
>
> _V_._P_. commentary, p. 54, Book III.

Unfortunately, there is no clear distinction at this stage between meaning in this sense and meaning in the sense of external objects or day-to-day substances. There was a big controversy running down the ages and extensively alluded to in the literature on philosophy of grammar between theorists who held that a word signifies necessarily an _individual_ or a class of them and others who held that it is always a universal _common_ _nature_ or a _quality_ which is designated by a word. The _Naiyāyikas_ tried to mediate by contending quite commonsensically that it is always certain individuals _qua_ bearers of certain word-intended properties that are denoted by words. Bhartṛhari, however, strives to maintain neutrality in this debate by holding that whether a substance or property--what is meant by a word--has only a conceptual being or intentional existence to begin with.

The notion of an individual substance (in the secondary non-transcendental sense) which he had is the widest possible one. It assigns _thinghood_ to everything, provided it is named. Thus, the notion has a superficial resemblance with Frege's notion of an object. Even a universal concept (which is usually contrasted to an individual substance) becomes a "substance" when it is brought under further characterising concepts. Wisdom is the substnace about which we assert "Wisdom is rare". "The concept horse" is a singular designator of an _object_.

In this way, "Whatever is referred to by pronominal words like 'that' and 'this' and which is purported to be distinguished by virtue of its being expressed by some name or other" is said to be a substance in this sense. However primitive such a definition might seem, it surely gives us the most liberal and all inclusive account of how we use the word "thing". This is the sense in which we can confer thinghood upon even a fictional object like Santa Claus once it has

become the focus of so many distinguishing remarks and intersubjectively available demonstrative thoughts (of children). When these vikalpas have nothing at all answering to them in the external world, we call them fictional, imaginary or unreal. Given these basic notions of a thing as the meaning of a noun-like expression let us see how Bhartṛhari tackles existence assertions and existence denials.

In V.P. 3/3 he discusses the distinction between existence in the sense of primary external reality (bāhyāsattā) and being, in the sense of secondary minimal reality (aupacārikīsattā) to which all posited designata of all namelike expressions are entitled. The criterion of external reality, however, is never made very clear. It must be distinct from the reality (pāramārthikīsattā) which attaches to the Transcendental Substance because external reality is earned by particular entities by coming into existence and lost by going out of it. That is why, the commentator sometimes interprets it as existing in the present time. There is some indication in the text of defining it in terms of self-identity, e.g. when birth of an item is defined as "gaining itself"; but I am not sure that Bhartṛhari would agree to render "That snake is not externally real" as "that snake is not self-identical".

His theory of verbal communication (as explained by Helārāja)--however psychologistic it may sound--throws some light on the secondary sense of existence which is the being appropriate to the vikalpas.

> In the beginning, the word and the meant entity stay undistinguished in the understanding of the speaker. We should not think that the intended meaning, i.e. the thought to be conveyed, is devoid of the structure or division, which is later displayed in the sentence, of the words, because it is intuited in the form of an inner talk...[25] Hence the sound which is uttered properly through the instrumentality of the [vocal and other facial] places and organs of speech [unlike inarticulate cooings or whistlings] is already designated by proper pause and repeatable patterns after the essential form of the meant entities. When heard by the hearers, it sparks off in them the same essential form and, thus, makes the meaning known.
>
> Linguistic communication consists in the mutual transmission of the intended purports of the speaker and the hearer. Meaning--the mental entity--leaves its original form in which it is buried in the intellect of the speaker and it is the word which carries it over and offers it [as its own form] to the hearer. And the hearer

> understands, according to his own subjective disposition, the content which is carved by the word-concepts.
>
> V.P., Book III/3/p. 265 (Commentary)

Major points in the above account are shrouded by metaphors. One thing, however, is certain: this is not a code-conception of language according to which words are just transmitters of meanings. Word and meaning, language and thought are conceived of as so essential to each other that often, in Bhartṛhari's idiom, they become indistinguishable. Thought is literally described as having "a body etched with language" (śabdakhacita vapu). It is perhaps mysticism to think of words and meanings as identical in the literal sense, but for all practical purposes this amounts to emphasising the fact that words do not merely transmit meanings; they are their essential vehicles and without them meanings have no independent existence (see Dummett, (1978), p. 7).

That the same content is grasped by both the hearer and the speaker is ensured by the same set of objective but not necessarily externally real vikalpas which are uniquely picked out by the particular words used. The kind of objectivity or existence which thought-posits enjoy is called minimal or metaphorical sense of existence, as contrasted with actual external existence. This notion of secondary existence reminds us of various western philosophical parallels, like Mackie's notion of minimal existence (Mackie 1976), Prior's wider notion of existence (the sense in which there is an a is true iff any statement of the form $\ulcorner \phi a \urcorner$ is meaningful (Prior 1976, p. 116)) or what Prior calls "Bertrand Drunk's" sense of being that

> ... belongs to every conceivable term, to every possible object of thought--in short to everything that can possibly occur in any proposition, true or false, and to all such propositions themselves. Being belongs to whatever can be counted... Numbers, the Homeric gods, relations, chimeras, and four dimensional spaces all have being... to mention anything is to show that it is.
>
> Russell (1903), p. 449

The need for such a projected, superimposed and extended variety of being has always been felt by philosophers who wish to take our singular denials of existence on their face value. Hence, Russell's remark:

> ... this distinction is essential if we are ever to deny the existence of anything.
>
> Ibid., p. 449

and that is precisely the use into which Bhartṛhari puts it.

> The usefulness of having a place for this realm of minimal existence is that we can put the object of negation in that realm.
>
> V.P., Commentary, III/SS-42.

With this distinction in mind, Helārāja explains our denials of external existence in the following fashion. The secondary existence which we have just elucidated is explicitly stated to be "common to reality and unreality" (bhāvābhāvasādhārana which unmistakably reminds us of Meinong's characterization of Aussersein as "beyond being and non-being"). What is denied is first posited as a minimally existent intentional object. This suspended objectivity is required to account for the reference of our used names in the context of existence questions and wonderments as to identity. In "Does a exist?" or "Is a the same as b?", "a" and "b" cannot refer to an externally existent object directly, without making the questions unaskable. They must refer to a vikalpa or a couple of them about which we can significantly ask "Is it also backed-up by external existence?" or "Are this and that vikalpas backed up by the same external entity?"[26]

Although the externally real object enters into our discourse and thought only as an object of understanding, i.e. as a vikalpa, that it also exists outside thought (when it does) is expressed by the verb "exists" in its present tense form. What we negate or exclude by predicating "exists" to a vikalpa is a kind of being which is merely dependent upon our cognition.[27]

Compare this analysis of the excluder use of "exist" with the following comment by Frege:

> If my intention is realised when I refer to something with the expression "That lime-tree" then the thought expressed by the sentence "That lime-tree is my idea" must obviously be negated.
>
> Strawson (ed.) (1967), p. 28

To the fact that "we play fast and loose" with the referential apparatus of our language, indifferently using expressions like "Mr. Pickwick", "Abraham Lincoln", "Mrs. Thatcher",

"The house I dreamt of", "My sister's expected baby", "The difference between you and me" and "The square root of 25" in the subject position of sentences meant to be true, three responses have been deemed possible by Frank Jackson (1980).

> The most ontologically extravagant is to admit all purported referents into our ontology: differences, possibilities, good chances, fictional characters and so on.
>
> An intermediate position is to construct a hierarchy of being or existence. All putative referents are allowed a minimal existence. But to avoid the shocking thesis that tables and differences exist equally, various degrees to which something may exist are distinguished.
>
> Finally, the most austere position is both to refuse to distinguish different kinds of existence and to refuse to admit all putative referents willy-nilly.

While the Nyāya position comes very close to the austere Quinean one,[28] it is not exactly clear where to place Bhartṛhari. In his theory of things in the widest sense, Bhartṛhari seems to be ontologically extravagant, but in making the distinction between primary and minimal existence, he seems to adopt the second, intermediate position. As to the Buddhists, who sometimes talk like Meinongians but are basically subjective idealists of a peculiar sceptical grain, they do not seem to believe that words can, at all, refer to realities.

4. GAṄGEŚA

The correct construal of the cognition displayed in the statement "The rabbit-horn does not exist" became an important issue once again in the "New Logic" initiated by Gaṅgeśa---the 14th century philosopher of logic and language--who wrote in the same tradition as Udayana, but used the new idiom of qualificands and qualifiers of cognitive contents, of limitors of relational and property abstracts, of absentees of absences, etc. with unsurpassed rigour in formulating exact definitions of key concepts in the theories of knowledge, inference and ontic categories.[29] Here I shall give only a very brief summary of Gaṅgeśa's comment on the problem.[30]

Absences are posited as real objects of valid cognitions expressed by true negative statements like:

> David Lewis is not in Oxford.

Since the cognitive content expressed by this statement is

distinct from that expressed by

> The author of <u>Counterfactuals</u> is not in the city where All Souls College is situated.

(although both are made true by the same state of affairs), we have to admit, over and above (a) the item cognized as the absentee (the person Lewis), (b) the location of the absence (Oxford), two further elements, viz. (c) a qualifier of the absentee--the property under which the particular item is grasped <u>as</u> the absentee and (d) a qualifier of the location---the description under which the location is grasped <u>as</u> such. Now, normally we can take as the qualifier of the absentee only a property which actually subsists in the absentee (in their language: The limitor of absenteehood should be co-located with the absenteehood) but is it not possible for us to miss a thing in a certain place (where it actually exists) because we are looking for it under the wrong sort of descriptions? Gaṅgeśa admits that this is possible if the description is <u>relational</u>. Thus, we can correctly cognise that the taste of the candy does not exist in the candy in the same manner as the wrapping paper exists on it, technically: there is an absence of sweetness (even in a sweet candy) of which the candy is the location, and the relation of <u>contact</u> is the limitor of the absenteehood of the sweet taste.

Can we now parallely construe the absence of rabbit-horns <u>everywhere</u> as the absence of which (any old) horn is the absentee, everything. e.g. a cow (typically an animal <u>with</u> horns) is the location, and <u>being a rabbit-part</u> is the limitor of absenteehood? Gaṅgeśa argues we cannot. His major reasons are the following.

First, not only do the (a) and (b) elements of the absence complex--the absentee and the location--have to be real items, the property which serves as the limitor of absenteehood has also to be located in the absentee. A property assumes the role of the limitor of an absenteehood only as a true characteristic of the relevant absentee--<u>not independently</u>. Since horns themselves are never known as rabbit-parts their absenteehood cannot be limited by the property of being rabbit-parts. Absence of completely fictitious items is not to be admitted also because its range would be absolute. The <u>Nyāya</u> can admit only absences of something in some--but not in other--places, or of something at some--but not other---times. Thus, universally unresisted absences of unreal objects cannot be recognised as objective. For one thing, every location can be correctly cognised as that of more than one such absence, resulting in a crisis of identity.

Second, knowledge of absence is always <u>knowledge that</u>, with a complex structure like the following:

For a piece of knowledge of the form "L has the absence of (A as f)"--the content would be (to use Matilal's symbol for cognitional qualification "Q(x,y)")[31]

Q(L, Q (Absence, Q (A, f)))

Absences cannot but be apprehended as absences of something, and that something must be a bona-fide qualified object. One cannot claim to be acquainted with a lack as directly as with a patch of colour--it has to be known as a lack of something. Since unqualified cognition of absences is impossible, the qualification of the absentee which enters into the structure of the content has to be correct, i.e. founded on real characterization rather than make-believe superimposition or error. The form of the cognition is always like this: The x which is ϕish is not there in this location. If the x in question is not really ϕish, even the cognition of its not being somewhere becomes incorrect, and, therefore unfit to make us aware of a genuine absence. If someone mistakenly thinks that x is ϕish, he will not perceive the absence of it in all places, whereas if the error is detected, the x which is ϕish will no longer appear as the thing that is absent. It may be possible for someone to believe that oviparous bats are absolutely non-existent, but if we take his formulation of the belief literally we must admit that his negative belief is as mistaken as the belief that there are such bats. What cannot be validly found in a certain manner of presentation cannot be validly missed either in such a manner.

Gaṅgeśa, therefore, interprets statements like "The rabbit-horn does not exist" as reflecting an absence to which rabbits are locations and horns, as horns, are absentees. If we nag at this interpretation that, after all, there is no locative case-marker after the word "rabbit" in the verbalisation of the cognition, Gaṅgeśa can very well justify his occasional mistrust of apparent linguistic forms by the following sort of example. To report the fact that I do not possess a car, I might humourously say "My car was never bought". It would be quite perverse to think that I thereby commit myself to an absolute absentee, viz. my car. Commitment to the order of designated objects is not a property of what I utter but a property of my interpretation of my utterance.

NOTES

1. Na hi ekam padam nirarthakam paśyāmah, prefactory commentary on Nyāya Sūtra, 3.1.1.

2. See Cartwright (1960).

3. See Ockham, Summa Logicae, II, chapters 12, 14.

4. Grammatically using a singular term, this statement is evidently better taken as general. That does not make any substantial difference to the Nyāya theory of meaning because all terms, singular or general are supposed to apply to a thing or to each of a set of things by virtue of a reason for application, usually a property.

5. See Chakrabarti (1982).

6. For a lucid exposition of this model see Kishore Chakrabarti, The Logic of Gotama.

7. Tasmād anupākhye vipakse hetor vyatirekanivrttau vā vyatireke vā sahrdayānāṃ mūkatā eva śaranam..., Vācaspati.

8. The section of Ātma Tattva Viveka where Udayana discusses this has been partly translated by Matilal (1970) and the present controversy has been discussed in the background of the Russell-Meinong debate in Matilal (1971) chapter 4. I have avoided repeating the points he has already made and disagree with his suggested branding of the Buddhist position as Meinongian. I have also gone a bit further into the original book where Udayana raises other points relevant to our purposes, of course drawing heavily on some of the insights Matilal's discussion contains.

9. I have translated from the text quoted in full in McDermott (1969). Her translation is often unreliable.

10. Ibid.

11. See Bhartṛhari, Vākyapadīya / Sambandha-Samuddeśa, verses 20-28.

12. The statement concerned being Asādhikā pratijñā. Helāraja, the commentator writes, Api tu etad-varjam pratijñāntarānām tattā.

13. All the way, both parties take 'the non-existent' as if

it were a referring expression, at least a putatively referring one.

14. See Plantinga (1974).

15. Routley (1980), p. 40.

16. Matilal (1971), pp. 141-144.

17. The Sanskrit words for them are Prasajya pratiṣedha and Paryudāsa respectively.

18. Vaktṛtvaviviktasya avastuno niyamena upalambhād āhosvid vastuviviktasya vaktṛtvasya anupalambhāt? See also Shaw (1974) for a discussion on these points.

19. Geach comments in Form and Existence (1955) that the change criterion of actuality suffers from the circularity that the change again has to be an actual change.

20. See Matilal (1971), p. 135.

21. Tathāca ananvitāh padārthā'eva anvitatayā parisphuranti iti viparītakhyātir evānuvartate

22. Śaśaviṣānādiśabdam uccārayataḥ kaścid abhiprāyo vṛtta iti tadviṣayo'sya vācya iti sugrahaḥ samayaḥ.

23. Udạyana compares such putative items of private reference to those of a dumb man's dreams (mūkasvapnavat in A.T.V., p. 77).

24. The apparently continuous circle of fire which is seen if a burning cinder is rotated very swiftly on a chain.

25. Antaḥ-saṃjalpa.

26. For a similar recent view see Bradford (1980), pp. 121-130.

27. Pratītimātraśaraṇā tu itarā vyudasyate.

28. Branding the Nyāya position as Russellian would be inaccurate. Existence is treated very much like a genuine property inasmuch as the sentence 'The jar exists' is taken as the paradigm example of a subject-predicate statement in their theory of meaning. Gaṅgeśa, for one, would maintain that even if unlocatables like 'being a rabbit-horn' are nonproperties, unnegatables like exis-

tence or knowability are properties. Here is one more point where he differs from Russell. He would not agree with Russell that 'There is no sort of point in a predicate which could not conceivably be false' (Russell 1918, p. 241).

Not distinguishing between properties and individuals rigidly, Gaṅgeśa argues that if being lacked by at least one individual is a necessary property of all properties then, it will itself be a universal property which is not lacked by any (relevant) individual (i.e. by any property). If on the other hand, it is itself lacked by some property, then some property is not lacked by at least one individual. Both ways therefore, we have to admit some universal property or other. The argument can be easily rebutted, but there is some important insight behind it. The notion of a property which takes sharability as central to it cannot demand contrastiveness as an indispensable feature of a property. If property is what can be exemplified by or can occur in something, what occurs in or is exemplified by everything has also to be regarded as a property (see Matilal 1968b), maybe a limiting case of a property.

29. For an introduction to the technicalities of Gaṅgeśa's idiom see Daniel Ingalls (1951).

30. A translation of the relevant Navya-Nyāya texts discussing this issue is to be found in Bhattacharya (1978).

31. Matilal (1968), p. 29. See also section 6.6, pp. 59-61 for discussion on the same problem.

REFERENCES

Bhattacharya, G.: 1978, Navya-Nyāya--Some Logical Problems in Historical Perspective, Delhi.

Bhartṛhari: 1974, Vākyapādiya, vols. I and III, ed. Raghunath Sharma, Varanasi.

Cartwright, R.L.: 1963, 'Negative Existentials', in C.E. Caton (ed.), Philosophy and Ordinary Language, University of Illinois Press.

Chakrabarti, A.: 1982, 'The Nyāya Proofs for the Existence of the Soul', Journal of Indian Philosophy, September 1982.

Chakrabarti, K.: 1978, The Logic of Gotama, University of Hawaii.

Dummett, M.: 1978, 'What Do We Know When We Know a Language?', lecture at Stockholm University, May 24, 1978.

Frege, G.: 1980, Translations from the Philosophical Writings of G. Frege, ed. Geach and Black, Blackwell.

Geach, P.T.: 1955, 'Form and Existence', P.A.S. 55.

Ingalls, D.H.H.: 1951, Materials for the Study of Navya-Nyāya Logic, Harvard.

McDermott, A.C.S.: 1970, An Eleventh-Century Buddhist Logic of "Exists", D. Reidel, Dordrecht.

Mackie, J.L.: 1976, 'The Riddle of Existence', symposium paper in P.A.S.S.V..

Matilal, B.K.: 1968, The Navya-Nyāya Doctrine of Negation, Harvard.

Matilal, B.K.: 1968b, 'Gaṅgeśa on the Concept of Universal Property', Philosophy East and West 18, no. 3.

Matilal, B.K.: 1971, Epistemology, Logic, and Grammar in Indian Philosophical Analysis, The Hague.

Ockham: 197 , Ockham's Theory of Proposition, translation of Summa Logicae II, by M.J. Loux, University of Notre Dame.

Plantinga, A.: 1974, The Nature of Necessity, Oxford University Press.

Prior, A.N.: 1976, 'Nonentities', in Papers in Logic and Ethics, Duckworth, London, pp. 109-121.

Routley, R.: 1980, Exploring Meinong's Jungle and Beyond, Australian National University.

Russell, B.: 1918, 'The Philosophy of Logical Atomism', in R.C. Marsh (ed.), Logic and Knowledge, London.

Shaw, J.L.: 1974, 'Empty Terms: The Nyāya and the Buddhists', Journal of Indian Philosophy 2, 332-343.

Strawson, P.F. (ed.): Philosophical Logic, Oxford.

Udayana: 1940, Ātmatattva Viveka, ed. D. Shastri, Benares.

J.L. Shaw

PROPER NAMES: CONTEMPORARY PHILOSOPHY AND THE NYĀYA

The aim of this paper is to present the Nyāya theories of proper names in light of the views of Frege, Russell, and Kripke. The first section of this paper will deal with the views of certain contemporary philosophers on proper names. Since the contemporary discussion is mainly confined to the description theory and the causal theory of proper names, I would like to discuss in this context the views of Frege, Russell, and Kripke. The second section will deal with the views of classical Nyāya philosophers on this topic. In this section I would like to discuss four different views of the Nyāya philosophers, and to compare their views with some contemporary philosophers. This discussion will take place in the light of the Nyāya theory of meaning. At the end of this section I would like to develop a theory of proper names, which, I think, can be developed within the Nyāya system. Moreover, this view may be claimed as a mean between the two other views of the Nyāya philosophers.

1.

It is claimed that Frege has not given a precise definition of a proper name. What he meant by the term 'proper name' can be explicated by his use of the term 'object'. An object is the referent of a proper name. Moreover, it is also claimed that his use of the term 'proper name' does not draw any distinction between a proper name and a definite description. In other words, all singular terms are considered as proper names. Dummett claimed that Frege has not given any precise characterization of 'proper names'. He said

> He (Frege) usually contented himself with using as a criterion the fact that an expression constituted a substantival phrase in the singular, governed by the definite article. He remained indifferent to the fact that this criterion would be inapplicable to those languages which lack a distinction of form between singular and plural, or to those even more numerous languages which lack a definite article: and equally indifferent to the fact that, even in those languages to which the criterion is applicable, it

B. K. Matilal and J. L. Shaw (eds.), Analytical Philosophy in Comparative Perspective, 327–372.

> is inexact in both directions.[1]

This is a valid comment if a criterion of a proper name is based on the syntactical structure of a particular language. But Frege, as Dummett also claimed, did not propose merely a syntactical criterion of a proper name. The criterion is based on Frege's use of the term 'object'. Dummett claimed that "...Frege's use of the ontological term 'object' is strictly correlative to his use of the linguistic term 'proper name'."[2] So Frege's characterization of a proper name depends on his use of the term 'object'. His use of the term 'object' cannot be explained without reference to his use of the term 'concept' or 'function'. Frege claimed:

> A concept (as I understand the word) is predicative. On the other hand, a name of an object, a proper name, is quite incapable of being used as a grammatical predicate.[3]

What Frege is saying is that a predicate expression refers to a concept or function, and a proper name refers to an object. From this it does not follow that a proper name cannot be part of a predicative expression. In the sentence 'The morning star is Venus', the predicate is not simply 'Venus', and the word 'is' is not merely the copula. The word 'is' means 'is no other than'. Hence the predicate is 'no other than Venus'. The latter expression refers to a concept. The object denoted by the subject expression 'The morning star' falls under the concept _no other than Venus_. In the sentence 'The morning star is a planet', the predicate expression is 'a planet' which refers to the concept _planet_, and the object denoted by the proper name 'The morning star' falls under the concept _planet_. But in the sentence 'The concept of _planet_ is acquired from experience', the subject term 'The concept of _planet_' does not refer to a concept, but to an object. Hence the concept of _planet_ in this context cannot be said to be a concept. This remark has puzzled several interpreters of Frege, and the literature on this topic has grown over the last three decades. Some of the comments of Frege are also responsible for the controversy. He has not made it explicit whether he is offering a linguistic criterion, or an ontological criterion, or an epistemic criterion, or a hint for drawing the distinction between a proper name and a predicative expression.

In his article 'On Concept and Object' he said:

> ...the singular definite article always indicates an object, whereas the indefinite article accompanies a concept-word.[4]

Again he said:

> ... I was not trying to give a definition, but only hints.[5]

If the criterion of Frege is just a linguistic criterion, then it is not universally valid. If it is just a hint, then he has not given any definite criterion for such a fundamental distinction between concept and object. It seems to me that a definite criterion can be formulated from the use he has made of these terms. Let us consider his examples:[6]

(a) There is at least one square root of 4.

(b) The concept _square_ _root_ _of_ _4_ is realized.

In both (a) and (b), he claimed, the same thought is expressed, but in (a) we are saying something about the concept _square_ _root_ _of_ _4_, and in (b) we are saying something about the object expressed by the proper name 'The concept _square_ _root_ _of_ _4_'. "This will be surprising only to somebody who fails to see that a thought can be split up in many ways, so that now one thing, now another, appears as subject or predicate."[7] This remark of Frege suggests that the object-concept distinction, at least in this case, depends on our manner of splitting up our thoughts. If this is considered as the criterion for the object-concept distinction, then it may be considered as an epistemic criterion. In the context of our discussion of Nyāya philosophy, we shall see whether Frege's distinction between object and concept corresponds to the Nyāya distinction between qualificand and qualifier.

Since both Frege and the Nyāya have introduced a sense theory of proper names, let us discuss Frege's arguments for the sense of a proper name. We can trace at least four different arguments for the acceptance of the sense of a proper name.

(1) According to Frege we cannot explain the difference between the statments 'a=a' and 'a=b', if 'a=b' is true, unless we admit that 'a' and 'b' have different senses. He claimed:

> If the sign 'a' differs from the sign 'b' only as an object (hereby its shape) but not by its role as a sign, that is to say, not in the manner in which it designates anything, then the cognitive significance of 'a=a' would be essentially the same as that of 'a=b', if 'a=b' is true.[8]

This remark shows that there is a fundamental difference

between these two types of identity statements. Moreover, he has claimed that 'a=a' is an analytic statement, whereas 'a=b' is a synthetic statement and "cannot always be justified in an a priori manner."[9]

(2) It is also claimed that we cannot identify the referent of a proper name unless we admit the sense of a proper name. Let us consider the term 'Aristotle'. Since the referent of this term cannot be presented to us, we can refer to him in terms of certain identifying descriptions such as 'the teacher of Alexander the Great'. Frege admits fluctuations in the criterion of identity for different people. That is to say, different people might identify Aristotle in terms of different identifying descriptions. Now the question is whether there is any need for an identifying description if the referent of a proper name is presented to us. A Fregean would still claim that there is a need for an identifying description if we want to distinguish the referents of different proper names. The referents of 'a' and 'b' cannot be distinguished unless we admit different identifying descriptions. The criteria for identity are usually expressed in terms of definite descriptions. The sense of a proper name usually provides us with such a criterion of identity. If we know the sense of a proper name, then we can identify the referent of it.

(3) Moreover, Frege has admitted the meaningfulness of expressions like 'the heavenly body which has the greatest distance from earth', and 'the series with the least convergence'. A nominatum for the first one cannot be guaranteed, and there is no nominatum or referent for the second one. Unless we admit senses for these expressions which are considered as proper names, they are to be regarded as meaningless expressions. According to Frege we can understand the meanings of these names even if they do not refer to any object. So the meaning of a proper name cannot be equated with its referent.

(4) Another argument for the sense theory can be construed from his remarks about the indirect discourse. In an indirect discourse, according to Frege, we do not refer to any customary nominata. He said:

> In indirect (oblique) discourse we speak of the sense, e.g., of the words of someone else. From this it becomes clear that also in indirect discourse words do not have their customary nominata; they here name what customarily would be their sense.[10]

Dummett does not consider this remark as an independent argument for the notion of sense. He claimed:

> ...it would be useless to offer any such explanation unless it had first been established that there is something which, in ordinary contexts, constitutes the sense of a term.[11]

This criticism is not applicable to our exposition of the sense of a proper name. We have established that in ordinary contexts there is a need for the sense of a proper name, and then we have claimed that the sense of a name, according to Frege, plays an important role in an indirect context. As we cannot explain the meaning of a sentence which contains an empty definite description unless we admit the sense of it, so we cannot explain the meaning of an indirect discourse in which a proper name occurs unless we admit the sense of it which is the referent in this context. Frege's aim was to assign meaning or sense to sentences like 'x believes that the heavenly body which has the greatest distance from the earth is round'. According to Frege in this case the subordinate clause has as its referent a proposition, not a truth-value.[12] If we admit a purely reference theory of proper names including the definite descriptions which are proper names on Frege's thesis, then it is difficult to give an analysis of oblique contexts. Frege's explanation in this respect is simpler than many other alternatives. So this may be considered as an additional argument for the sense of a name.

Now let us discuss the nature of the sense of a proper name. Frege has made the following remarks about the sense of a proper name.

(a) It is a mode of presentation of the object which is the referent of a proper name, or it is the manner in which a name designates its referent or the manner in which an object is presented to us.[13]

(b) The sense of a proper name contains the manner and context of presentation.[14]

(c) "The sense of a proper name is grasped by everyone who knows the language..."[15]

(d) It illuminates the nominatum or referent if there is any.[16]

(e) It is also claimed that the sense of a proper name belongs to the nominatum if there is one.[17]

(f) The step from the sense of a proper name to its nominatum is determinate.[18] In other words, the sense of a proper name determines its nominatum.

All these different ways of describing the sense of a proper name do not seem to be equivalent. From (a) it does not follow that the sense of a proper name has an ontological status or it belongs to the referent. (b) suggests that the sense contains an epistemic element, and it remains an open question whether it is identical with the manner of

presentation. Moreover, both (a) and (b) are consistent with the fluctuations in the sense of a proper name. As a matter of fact Frege himself has admitted a fluctuation in the sense of a proper name. (c) suggests that the sense of a proper name is a part of our knowledge of language. Knowing a language is an activity in which human beings are engaged, and this includes knowing the sense of a proper name. (c) by itself does not entail the other characterizations of the sense of a proper name. It simply emphasizes one of the conditions for grasping the sense of a name. (d) emphasizes one of the functions of a proper name. It is compared with light. As light illuminates the things of a particular place, so the sense illuminates the referent if there is any. This metaphor also suggests that as light does not need anything else to reveal itself, so the sense does not need another sense to reveal itself. As a light is an instrument for knowing an object, so the sense is an instrument for knowing the referent of an expression. (e) suggests something different from the rest of the characterizations of the sense of a proper name. It seems to me that this characterization should not be taken literally. If it is taken literally, then we have to say that as a book belongs to me or a shade of blue color belongs to this book, so the sense belongs to its referent if there is any. But this does not seem to be a plausible explanation of sense if we take into account the rest of the things he has said about sense. What he meant is that corresponding to a sense there is a property which belongs to the referent of a proper name if there is any. If the sense of the name 'Aristotle' is expressed by 'the teacher of Alexander the Great', then there is a property of being the teacher of Alexander the Great which belongs to Aristotle. (f) emphasizes how the sense of a proper name leads to its referent. It is a path which leads us to the referent if there is any. This characterization emphasizes how a linguistic expression is related to its referent. It is not related directly to its referent, but <u>via</u> sense.

In addition to the above descriptions of sense, Frege has distinguished sense from ideas on the one hand, and physical objects on the other. It is neither subjective like ideas or images, nor is it objective like physical objects. It is in between these two extremes. It is objective in the sense that many people can have the same sense, which is not possible with respect to our ideas or images. Since it is not like physical objects it cannot be located in some space and time, assuming that only physical objects are in space and time. It is neither psychological nor physical, but an ontological entity which can be grasped by human beings or by those who use language for communicating thoughts. In the context of our discussion of the Nyāya view we shall see how far the Nyāya concept of the limitor corresponds to Frege's

conception of sense. Moreover, in the context of Kripke's theory of proper names we shall discuss whether there are different types of sense of a proper name according to Frege.

Now let us discuss Russell's conception of a proper name. According to Russell, unlike Frege, there is a sharp distinction between a logically proper name on the one hand, and an ordinary proper name or a definite description on the other. Since an ordinary proper name like 'Socrates' is a set of definite descriptions in disguise, we can contrast a logically proper name with a definite description. This discussion would also be useful for understanding the Nyāya distinction between two types of singular terms used to designate the particular objects.

Russell has explained the distinction between a logically proper name and all other symbols in the following way:

> A name is a simple symbol whose meaning is something that can only occur as subject, i.e., something of the kind that, ... , we defined as an "individual" or a "particular". And a "simple" symbol is one which has no parts that are symbols. Thus "Scott" is a simple symbol, because, though it has parts (namely, separate letters), these parts are not symbols. On the other hand, "the author of Waverley" is not a simple symbol, because the separate words that compose the phrase are parts which are symbols.[19]

Again he said:

> We have, then two things to compare:
> (1) a name, which is a simple symbol, directly designating an individual which is its meaning, and having this meaning in its own right, independently of the meanings of all other words;
> (2) a description, which consists of several words, whose meanings are already fixed, and from which results whatever is to be taken as the "meaning" of the description.[20]

From the above remarks it follows that a name is a simple symbol, and the object it denotes can only occur as subject. In other words, a name in a sentence is to be treated as a subject-expression, and it denotes an individual or a particular which is its meaning.

Moreover, Russell has contrasted a logically proper name with a definite description. A proper name is a complete symbol in the sense that it has a meaning in isolation, while a definite description has no meaning in isolation. It has a

meaning in <u>use</u> or in the context of a proposition. The proposition 'Scott is mortal', where 'Scott' is a logically proper name, and the proposition 'The author of <u>Waverley</u> is mortal' are not of the same form. With respect to these two types of proposition Russell has said:

> You think that they are both simple propositions attributing a predicate to a subject. That is an entire delusion: one of them is (or rather might be) and one of them is not. These things, "the author of <u>Waverley</u>", which I call incomplete symbols, are things that have absolutely no meaning whatsoever in isolation but merely acquire a meaning in a context. "Scott" taken as a name has a meaning all by itself.[21]

Furthermore, a logically proper name cannot be eliminated from the sentence in which it occurs, but a definite description can be eliminated from the sentence in which it occurs. This follows from Russell's theory of definite descriptions.

The sole function of a name, according to Russell, is to denote a particular. He said:

> A name can just name a particular, or, if it does not, it is not a name at all, it is a noise.[22]

Unlike Frege, Russell claimed that a sentence is not a name. The relation of a name to the thing named is totally different from that of a proposition to fact. A proposition which is "a sentence in the indicative, a sentence asserting something, not questioning or commanding or wishing,"[23] is not a name for a fact; it is not a name for anything.

Now let us discuss the nature of a denotatum of a logically proper name. According to Russell the particulars which are denotata of proper names are momentary things like patches of colors, or sounds. The entities like Piccadilly, Rumania, and Socrates are not considered as particulars in the strict sense. The denotation of the word 'Piccadilly' is a series of classes of material entities.[24] The logical status of Piccadilly is bound up with the logical status of series and classes. Since according to Russell a series or a class is a logical fiction, an object like Piccadilly or Socrates is also a logical fiction. He said:

> As you know, I believe that series and classes are of the nature of logical fictions: therefore that thesis, if it can be maintained, will dissolve Piccadilly into a fiction. Exactly similar remarks will apply to other instances:

> Rumania, Twelfth Night, and Socrates.[25]

But, according to Russell, a particular which is the denotatum of a logically proper name is not a logical fiction. It occurs as a term of a relation in an atomic fact.

As regards the mode of knowledge of a particular Russell claimed:

> A name, ... , can only be applied to a particular with which the speaker is acquainted, because you cannot name anything you are not acquainted with.[26]

From this remark it follows that we cannot name an object or a particular unless we are acquainted with it. This condition does not apply to a description or an ordinary name. Since we are not acquainted with the person Socrates, we cannot name him by using the name 'Socrates'. The name 'Socrates' is simply an abbreviated description. The understanding of a descriptive expression or predicate-expression, according to Russell, involves an act of mind different from the act which is involved in understanding a proper name. He said:

> To understand a name you must be acquainted with the particular of which it is a name, and you must know that it is a name of that particular. You do not, that is to say, have any suggestion of the form of a proposition, whereas in understanding a predicate you do. To understand "red" for instance, is to understand what is meant by saying that a thing is red.[27]

This remark suggests that a logically proper name is a complete expression, while other expressions are incomplete in the sense that the understanding of any expression other than a logically proper name involves understanding a proposition in which it occurs. The above discussion of Russell may be summarized in the following way:

(1) A proper name is a simple symbol, and it cannot be eliminated from a sentence in which it occurs.

(2) A proper name can only occur as a subject in a proposition.

(3) A proper name is a name for a particular.

(4) The meaning of a proper name is its referent or the particular it denotes.

(5) We cannot know the meaning of a proper name unless we are acquainted with its referent or the particular it denotes.

(6) The understanding of the meaning of a proper name does not involve understanding a proposition in which it

occurs.

The first two theses may be taken as linguistic or syntactic characterization of a proper name, (3) and (4) as semantic characterization, and the last two as epistemic characterization of a proper name.

Now let us discuss the view of Kripke who has criticized the views of philosophers like Frege, Russell, Strawson, Searle and Wittgenstein. It is claimed that he has introduced a new theory of proper names, which has not been introduced by any of his predecessors.

Kripke claimed that Frege and Russell have equated the meaning or the sense of a proper name with the sense of a definite description. With respect to this view he said: "I think it is pretty certain that the view of Frege and Russell is false."[28] According to him many philosophers who claim to reject the theory of Frege-Russell "have used the notion of cluster concept."[29] On Kripke's interpretation Wittgenstein has equated the meaning of a proper name with that of a family of descriptions. In order to substantiate this view he has quoted paragraph 79 from Wittgenstein's <u>Philosophical Investigations</u>. It is doubtful whether we should attribute such a theory of proper names to Wittgenstein. It seems to me that Wittgenstein is explaining certain uses of proper names in our ordinary language. Perhaps there is no one theory for explaining the different uses of a proper name according to Wittgenstein.

It is also claimed that according to the view of Searle or Strawson "the referent of a name is determined not by a single description but by some cluster or family."[30] Kripke has formulated the description theory of proper names in the following way:

> (1) To every name or designating expression 'X', there corresponds a cluster of properties, namely the family of those properties <u>Ø</u> such that <u>A</u> believes '<u>ØX</u>'.
>
> (2) One of the properties, or some conjointly, are believed by <u>A</u> to pick out some individual uniquely.
>
> (3) If most, or a weighted most, of the <u>Ø</u>'s are satisfied by one unique object <u>y</u>, then <u>y</u> is the referent of '<u>X</u>'.
>
> (4) If the vote yields no unique object, '<u>X</u>' does not refer.
>
> (5) The statement, 'If <u>X</u> exists, then <u>X</u> has most of the <u>Ø</u>'s' is known <u>a priori</u> by the

speaker.

> (6) The statement, 'If X exists, then X has most of the Ø's' expresses a necessary truth (in the idiolect of the speaker).[31]

According to Kripke thesis (1) is a definition of the cluster theory of names, and theses (2)-(6) are false. Thesis (2) is false, because we can say that Feynman is a famous physicist without attributing anything else to Feynman. Here the description is not satisfied by someone uniquely. According to Kripke the name 'Feynman' refers to Feynman even if we fail to associate any definite description with the name.

With respect to thesis (3) he said that even if most of the Ø's are satisfied by a unique object y, y might not be the referent of the name 'X'. Kripke said that we usually associate the definite description 'the man who proved the incompleteness of arithmetic' with the name 'Gödel'. Let us suppose that Gödel was not the author of this theorem, but someone else called 'Schmidt' was the author of the theorem. Kripke said:

> So, since the man who discovered the incompleteness of arithmetic is in fact Schmidt, we, when we talk about 'Gödel', are in fact always referring to Schmidt. But it seems to me that we are not.[32]

From this remark of Kripke it follows that even if most of the Ø's are satisfied by a unique object, that object might not be the referent of a name.

With respect to thesis (4) he said that even if nothing satisfies a set of descriptions associated with a name, it does not follow that a name does not refer to anything. With respect to the name 'Jonah' Kripke said:

> Biblical scholars, ..., think Jonah really existed. It isn't because they think someone ever was swallowed by a big fish or even went to Nineveh to preach. These conditions may be true of no one whatsoever and yet the name 'Jonah' really has a referent.[33]

From this remark of Kripke it follows that the name 'Jonah' refers to Jonah even if it does not satisfy any description uniquely.

Thesis (5) is also false. Kripke said:

> ...even in a case when (3) and (4) happen to be true, a typical speaker hardly knows a priori that they are, as required by the theory.[34]

So, according to Kripke, the fact that most of the descriptions are uniquely satisfied by an object is not known a priori by the speaker. For example, that Gödel satisfies the description 'the man who proved the incompleteness of arithmetic' is an empirical knowledge.

Thesis (6) is false, because it is not a necessary truth that if X exists, then X has most of the Ø's. Let us consider the name 'Aristotle'. The properties like being the teacher of Alexander the Great and being the author of such and such books are attributed to Aristotle. According to Kripke, even if Aristotle exists, it does not follow that he satisfies most of the properties usually ascribed to him in all possible worlds. Hence we cannot claim that it is a necessary truth that if X exists, then X has most of the Ø's.

According to Kripke the above theories of names are false, because the supporters of those theories did not realize that proper names are rigid designators. A name is a rigid designator if, in every possible world, it designates the same object. A designator is non-rigid if it does not designate the same object in all possible worlds. Most of the definite descriptions are non-rigid designators. The expression 'the teacher of Alexander the Great' is not a rigid designator, because in different possible worlds different persons might satisfy this description. But the name 'Aristotle' designates the same person in all possible worlds even if all the properties ascribed to him are not satisfied by him.

Kripke has also claimed that a modal context makes the distinction between a proper name and a definite description transparent. A definite description is capable of scope distinctions in modal contexts, while a proper name is not. Let us consider the following propositions.

(1) Aristotle is the teacher of Alexander the Great.

(2) The teacher of Alexander the Great is the teacher of Alexander the Great.

(3) Aristotle is Aristotle.

Since the property of being the teacher of Alexander the Great is not an essential property of Aristotle, we can assert 'Aristotle might not have been the teacher of Alexander the Great'. The latter proposition can be symbolized in the following way:

(1') $(\exists y)(y{=}\text{Aristotle}.\ \Diamond\ y \neq (\iota x)Fx)$, where

'$(\iota x)Fx$' means 'the teacher of Alexander the

Great'.

That is to say, in some possible world the same person is designated by the name 'Aristotle', but he is not the teacher of Alexander the Great. So in this world the expression 'the teacher of Alexander the Great' is used referentially to fix the referent of the term 'Aristotle'. Hence the definite description 'the teacher of Alexander the Great' which is not a rigid designator cannot be equated with the name 'Aristotle' which is a rigid designator.

Similarly, we can assert that the teacher of Alexander the Great might not have taught Alexander the Great. But we cannot assert that Aristotle might not have been Aristotle. The former proposition can be symbolized in the following way:

(2') $(\exists \underline{x})\ ((T\underline{x}.(\underline{y})(\underline{Ty} \supset \underline{x=y})).\ \Diamond \sim \underline{Tx})$, <u>where</u>

'<u>Tx</u>' means '<u>x</u> taught Alexander the Great'.

The proposition 'The teacher of Alexander the Great might not have taught Alexander the Great' is not identical with the proposition 'It is possible that the teacher of Alexander the Great did not teach Alexander the Great'. The latter proposition can be symbolized in the following way:

(2'') $\Diamond\ (\exists x)\ ((Tx\ .\ (y)(Ty \supset x=y)).\ \sim Tx)$.

(2') is true, but not (2''). (2) is true in this world, and (2') is true in some possible world; but (2'') is not true in any possible world. According to Kripke the truth-value of the proposition 'The teacher of Alexander the Great might not have been the teacher of Alexander the Great' would change if we substitute proper names for definite descriptions. The proposition 'Aristotle might not have been Aristotle' is false in every possible world. This sentence is not capable of scope distinctions according to Kripke. It can be symbolized in the following way:

(3') $\Diamond$ (Aristotle $\neq$ Aristotle.)[35]

From the above discussion it follows that according to Kripke a proper name designates the same object in every possible world, but a definite description of the sort mentioned above does not designate the same object in every possible world. When we say that Aristotle is the the teacher of Alexander the Great, we are not, according to Kripke, describing the sense of the name 'Aristotle' in terms of the sense of 'the teacher of Alexander the Great'. We are simply fixing the reference of the name 'Aristotle' in terms of a definite description used referentially. In another possible world the same

description might not pick out the same person.

From the above distinction between a proper name and a definite description we should not conclude that all definite descriptions are referential, and hence non-rigid. Kripke claimed that a definite description which corresponds to an essential property or a structural property of an object is a rigid designator. For this reason a definite description like 'the metal with the atomic number 79' is considered as a rigid designator. Hence in addition to the traditional classification of singular terms into proper names and definite descriptions Kripke has introduced the classification of singular terms into rigid and non-rigid designators.

As to the nature of the referent of a proper name, Kripke said that it is a particular which is referred to by a proper name, but it cannot be equated with 'a bundle of qualities'. Moreover, it cannot be considered as something which is behind these qualities.

Now let us discuss the causal or historical explanation theory of proper names, which is attributed to Kripke. It is doubtful whether Kripke is proposing another theory of proper names in the sense in which the word 'theory' is used in other contexts. In course of his discussion of the cluster concept theory of names he said:

> It's wrong. You may suspect me of proposing another theory in its place; but I hope not, because I'm sure it's wrong too if it is a theory.[36]

In spite of this remark he has discussed what most of us do or how most of us fix the referent of a proper name, or how most of us refer to the referent of a proper name. He said:

> Someone, let us say, a baby, is born; his parents call him by a certain name. They talk about him to their friends. Other people meet him. Through various sorts of talk the name is spread from link to link as if by a chain. A speaker who is on the far end of this chain, who has heard about, say Richard Feynman, in the market place or elsewhere, may be referring to Richard Feynman even though he can't remember from whom he first heard of Feynman or from whom he ever heard of Feynman. He knows that Feynman was a famous physicist. A certain passage of communication reaching ultimately to the man himself does reach the speaker. He then is referring to Feynman even though he can't identify him uniquely.[37]

From this remark of Kripke it follows that a name, say 'Feynman', refers to the same person even if its subsequent users cannot identify that person or even if they cannot remember from whom they have heard of the name 'Feynman'. Again he said:

> On our view, it is not how the speaker thinks he got the reference, but the actual chain of communication, which is relevant.[38]

Kripke repeatedly emphasizes the chain of communication for successful reference.

> ...what is true is that it's in virtue of our connection with other speakers in the community, going back to the referent himself, that we refer to a certain man.[39]
>
> In general our reference depends not just on what we think ourselves, but on other people in the community, the history of how the name reached one, and things like that. It is by following such a history that one gets to the reference.[40]

In addition to the community of the speakers and the actual chain of communication or the history of the name, Kripke has introduced the identity of intention of the speakers for a successful reference. He said:

> When the name is 'passed from link to link', the receiver of the name must, I think, intend when he learns it to use it with the same reference as the man from whom he heard it.[41]

From the above remarks of Kripke it follows that he is emphasizing three aspects. (1) The initial baptism which fixes the reference either by ostension or by a description. Since only structural properties are considered as essential ones, the descriptions used in this context are usually referential and hence non-rigid. (2) The causal chain or subsequent users of a name, who form the community of speakers of a particular name. This point emphasizes the social aspect of the use of a name. (3) The identity of intention of speakers, which serves as the cement for the chain of communication. If these conditions are fulfilled, then the name or the user of a name can refer to the person or object which was baptized initially by a particular person. But if there is a change in the intention of the speakers, then we get cases like the name 'Santa Claus'. The name 'Santa Claus', according to Kripke, was originally used to refer to a certain saint. Since there

was some change in the intention of subsequent users, this name as used by children does not refer to that saint.

Now let us discuss whether the objections Kripke has raised against the sense theory of Frege are valid, and whether Kripke's causal chain theory can be considered as a substitute for Frege's sense theory of proper names.

From our discussion of Frege it follows that he has used the word 'sense' in different senses. If we take some of the uses of the word 'sense', then we cannot claim that according to Frege the sense of a proper name is always expressible by a definite description. In other words, it cannot be said that being expressible by a definite description is a necessary condition for having sense. Since the sense of a name is a mode of presentation, it might not always be expressible by a definite description. When the sense is adequate for determining the referent of a name, it is expressible by a definite description. Frege's use of the term 'sense' is essentially epistemic, and it explains what happens when one uses a name or one understands a name. But Kripke has emphasized only the social or the historical aspect of the use of a name. According to Frege the user of a name refers to an object *via* its sense, while according to Kripke a name refers to an object *via* the causal chain and the identity of the intention of the speakers. According to Frege sometimes a sense can be expressed by a definite description, and sometimes by an indefinite description. The sense of the name 'Feynman' might sometimes be expressed by the term, 'a famous physicist'. Hence this example of Kripke does not constitute a counterexample to Frege, or a refutation of Frege's theory of sense. The sense expressed by 'a famous physicist' is not adequate for determining or recognizing the referent of a name. But this cannot establish the thesis that an indeterminate sense is not a sense at all. An indeterminate sense which is expressible by an indefinite description might be considered as an incomplete sense of a name. From some remarks of Frege one can develop three different types of sense of a name which can be known by the user of a name.

(1) Incomplete sense: an incomplete sense of a name is expressible by an indefinite description and it is not sufficient for determining the referent uniquely.

(2) Partially complete sense: this type of sense would be sufficient for a particular user of a name to identify the referent in a particular world or possible world. Hence it can be expressed by a definite description.

(3) Semi-complete sense: this type of sense would give us a necessary and sufficient condition for the identity of the referent in all possible worlds. The epistemic counterpart of an essential and a unique property of the referent would be a sense of this type. If we can cognize an object under the mode of an essential and a unique property, then we

can identify it in any possible world. The epistemic counterpart of Kripke's description of an object in terms of its structural property would be a sense of this type.

The term 'complete sense', according to Frege, is reserved for all possible modes of presentation of an object. But a finite human being cannot grasp all possible modes of presentation of an object. Frege has admitted this point in the following passage:

> A complete knowledge of the nominatum would require that we could tell immediately in the case of any given sense whether it belongs to the nominatum. This we shall never be able to do.[42]

If we try to understand Frege's theory in the light of different senses of the word 'sense' and in light of different types of sense, then some of the remarks or counterexamples of Kripke are false or wide of the mark. Kripke's example 'Feynman was a famous physicist' does not constitute a counterexample to Frege's theory of sense. It simply shows that the mode of presentation of Feynman is an incomplete sense.

Secondly, Kripke's method of fixing the referent of a name in terms of a referential definite description does not constitute a counterexample to Frege's theory of sense. A referential definite description which is satisfied by an object would express a partial sense of a name. Moreover, Frege has admitted a fluctuation in the sense of a name depending on the context and the speaker, as well as the cognitive value of this fluctuation. He said:

> In the case of genuinely proper names like 'Aristotle' opinions as regards their sense may diverge. As such may, e.g., be suggested: Plato's disciple and the teacher of Alexander the Great. Whoever accepts this sense will interpret the meaning of the statement "Aristotle was born in Stagira" differently from one who interpreted the sense of 'Aristotle' as the Stagirite teacher of Alexander the Great. As long as the nominatum remains the same, these fluctuations in sense are tolerable.[43]

From this remark of Frege it follows that the cognitive value or the meaning of a sentence depends on how we fix the referent of a name. So far as ordinary language is concerned Frege has admitted different ways of fixing the referent of a name. In a demonstrative science, according to Frege, these fluctuations should be avoided, and in a perfect language there should not be this type of fluctuation in the sense of a

proper name. Hence the different senses of the name 'Aristotle' which is part of the ordinary language are used to fix the referent of it. In this world the referent of the name 'Aristotle' can be fixed either in terms of 'the disciple of Plato' or in terms of 'the teacher of Alexander the Great' or 'the Stagirite teacher of Alexander the Great'.

Thirdly, if the referent of a name is fixed in terms of a structural property which is an essential and a unique property of an object, then also it does not constitute a counterexample to Frege. The mode of presentation which determines the referent in all possible worlds would be an epistemic counterpart of an essential and a unique property of the referent. A definite description of this type would be a rigid designator according to Kripke, and the mode of presentation of this type would be a semi-complete sense of Frege.

Fourthly, Kripke's causal or historical explanation theory of proper names cannot be taken either as a substitute for Frege's theory of sense or as a falsifying condition for Frege's theory of proper names. Frege is mainly concerned with the question what happens at the level of understanding when we understand a sentence in which a proper name occurs, and how a user of a name can identify the referent of a proper name. He is also concerned with the question what constitutes the knowing of a language. He said:

> The sense of a proper name is grasped by everyone who knows the language....[44]

Kripke, on the other hand, is concerned with the question how a name is actually linked with its referent. His answer in terms of a causal or historical explanation is not concerned with the epistemic questions of Frege. So Kripke is concerned with the question how a name is linked up with its referent, and not with the question how a user can identify the referent of a proper name. According to Kripke since a proper name is a rigid designator, it refers to the same object in every possible world even if a user of it cannot identify the referent in any possible world. Since Kripke and Frege are concerned with different questions, Kripke's theory cannot be claimed to be a substitute for Frege's sense theory which is an epistemic theory of proper names.

Fifthly, Kripke in his causal theory of names emphasizes the importance of the identity of intention of the users, which serves as the cement or the link in this causal chain. According to him the receiver of a name must intend to use it with the same reference as the person from whom he has received the name. This raises the question whether we know the intention of our previous users of a name. Kripke's theory can only work if we had a complete history of all the names used in our language and we could also determine the

identity of the intention of different users of a name. As we do not have such a history of proper names, his view does not seem to be a plausible theory of names. Moreover, since there is no history of all the names used in our language, sometimes we do not know whether such and such a person existed.[45]

From these criticisms of Kripke one should not conclude that we are not in favor of emphasizing the social aspect of the use of a name. I am simply pointing out certain difficulties if this theory of Kripke is taken as a substitute for Frege's sense theory of names or if it is taken as the most satisfactory explanation for what most of us do.[46] It seems to me that a satisfactory theory of proper names would emphasize both the epistemic and the social aspects of the different uses of a proper name in our ordinary language. In the next section we shall see that the Nyāya view emphasizes both the name giving ceremony and the intentional aspect of Kripke's theory, and the sense theory of Frege.

2.

In this section I would like to discuss the Nyāya theories of proper names in the light of our above discussion so that we can compare the Nyāya theories with the views of some contemporary philosophers. Before introducing the Nyāya discussion of proper names I would like to mention the Nyāya concept of meaning and cognition.

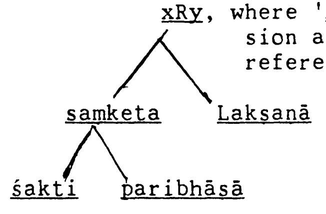
The Nyāya has admitted a conventional relation between an expression and what is referred to by that expression. This relation is called '<u>vṛtti</u>'. There are two types of this relation depending on whether an expression refers to the primary referent or the secondary referent of it. The usual referent or the referents of an expression are primary referents. For example, all individual cows are the primary referents of the word 'cow'. Similarly, the person John is the primary referent of the word 'John'. In the context of metaphor, according to the Nyāya, an expression does not refer to its usual referent, but it does refer to something which is related to its primary referent, and this secondary referent is the intended object of the speaker or the writer. For example, in the sentence 'The village is on the Ganges', the primary referent of the word 'Ganges' is the river Ganges, but the secondary referent is the bank of the river Ganges. In this sentence we should take the secondary referent of the word 'Ganges'.

The relation between an expression and its primary referent is called '<u>saṃketa</u>' ('primary relation'), and the relation between an expression and its secondary referent is called '<u>Lakṣaṇā</u>' ('secondary relation'). Again <u>saṃketa</u> or the primary relation is of two types depending on whether the relation is set up by the will or the intention of a

particular human being or by the will of God. Here the will of God is postulated simply to explain the conventions whose origins are not known to us. But in both cases the relation between an expression and its referent is due to the intention of the name-giver, and hence it is conventional. These two types of convention are called '*śakti*' (i.e., 'ancient convention'), and '*paribhāṣā*' (i.e., 'modern convention'). The relation between an expression such as 'cow' amd its referents is called '*śakti*', because we do not know when this convention was set up. But the relation between a proper name or a technical term coined by a scientist or a grammarian and its referent is called '*paribhāṣā*'.

Since the Nyāya has emphasized the direction of a relation, the relation of an expression to its referent is not the same as the relation of a referent to an expression. If the relation is set up by an intention of the form "Let such and such object be cognized by such and such expression",[47] then this relation is called '*arthaviśeṣyaka*', i.e., the object or the referent is the first term of the meaning-relation whose second term is the expression. But if the relation between an expression and its referent is set up by a convention (i.e., intention of the name-giver) of the form "Let such and such expression generate the cognition of such and such object",[48] then it is called '*padaviśeṣyaka*', i.e., an expression is the first term of the meaning-relation whose second term is the referent. The Nyāya classification of the relation between an expression and its referent may be represented by the following diagram:

xRy, where '*x*' stands for an expression and '*y*' stands for the referent of '*x*'.



Each of the terminal nodes can be classified further depending on the direction of the relation between an expression and its referent or referents.

From the above discussion it follows that the Nyāya emphasizes the conventional nature of the relation between an expression and its referent or referents, and the distinction between the two types of primary relation depending on whether the convention is ancient or modern.[49] The Nyāya use of the term '*śakti*' is to be interpreted as an ancient conventional relation between an expression and its referent or referents, and the term '*paribhāṣā*' is to be considered as a modern conventional relation between an expression and its referent

or referents.[50]

Now let us introduce the Nyāya discussion of the meaning of a common noun such as the word 'cow'. In this context I would like to introduce the neutral word 'meaning-complex' which refers to the second member of a meaning-relation whose first member is an expression. The Nyāya deals with the question whether the meaning-complex includes the particulars to which an expression applies or the universal of which the particulars are instances, or the configuration (ākrti) of the particulars. According to the old Nyāya the meaning-complex includes the universal or the class-character (jāti), the configuration (ākrti) of the particulars, and the particulars.[51] So the meaning-complex of the word 'cow' includes particular cows or the cow-individuals, the class-character, i.e., cowness, and the configuration (ākrti) of particular cows. According to the old Nyāya it is one meaning-relation which relates the word 'cow' to the meaning-complex which includes three types of entities.[52]

Now the question is how the expression 'particular cows or cow-individuals' and 'configuration of particular cows' are to be understood. By the word 'cow-individuals' we should understand particular cows without their qualities and relations. If the cow, say a, is white, and the cow, say b, is black, then by the word 'cow-individuals' we should understand a and b without their colors. Similarly, a and b are to be taken without their other properties except properties like being this or that individual. Even the class-character is excluded from a and b. The expression 'configuration (ākrti) of cow-individuals' refers to the parts of the cow-individuals, which are related in a particular manner in a particular cow. Let us consider the particular cows, a and b. Cow a has parts such as a tail, a head, four legs, and a body, which are related in a particular way. Similarly, cow b has the same type of parts which are related in the same way. Symbolically, the parts of a and b can be represented in the following way:

(1) $R\ (a_1, a_2, a_3, a_4)$,

(2) $R\ (b_1, b_2, b_3, b_4)$,

where the subscripts represent the different parts of cow a and b.

Now by the word 'configuration of cow-individuals' we should not mean the conjuntion of (1) and (2), but the generic features of (1) and (2), which can be represented in the following way:

(3) R $(\alpha, \beta, \gamma, \delta)$,

where 'α', 'β', 'γ', and 'δ' represent the <u>types</u> of parts of different cow-individuals.[53]

But the Navya-Nyāya has raised several objections against the view which includes the configuration of particulars within the meaning-complex of an expression.

First of all, this view cannot give a similar account of the meanings of expressions for qualities, actions, and other properties.[54] Let us consider the meaning of the word 'brown'. The meaning-complex of this expression includes only the particular brown colors and the universal <u>brownness</u>. Since a color does not have a configuration (<u>ākṛti</u>), the question whether it is a member of the meaning-complex does not arise.

Secondly, even in the case of a substance-referring expression it is not necessary to include the configuration within the meaning-complex of it. What we need in addition to the set of particular objects to which a word applies is a limitor (<u>avacchedaka</u>) which will <u>determine</u> the referents uniquely. The followers of the Navya-Nyāya claimed that with respect to the word 'cow' the universal cowness is sufficient to determine the set of particular cows. Since cowness is present in cows only, it will determine the set of cows uniquely. Hence we do not need the configuration of particular cows in addition to the universal cowness to determine the referents of the term 'cow'.

Thirdly, it cannot be said that configuration should be included in the meaning-complex of an expression instead of a class-character (<u>jāti</u>). Since a class-character is an atomic property and a configuration is a molecular property, on the ground of simplicity the class-character should be retained in the meaning-complex of an expression.[55] Moreover, according to Jagadīśa, a Navya-Naiyāyika, the term '<u>ākṛti</u>' ('configuration') mentioned in the <u>Nyāyasūtra</u> should not be interpreted as the form or configuration of an object. He claimed that the word '<u>ākṛti</u>' taken in its etymological sense means 'that which determines the form of an object'. Hence it is to be taken as the relation between the class-character and the individuals which are instances of the class-character.[56] In the case of the word 'cow' the meaning-complex includes the individual cows, the universal cowness, and the relation of inherence which relates the latter to the former. Hence the meaning-complex of the word 'cow' is an ordered set of three elements. It can be expressed in the following way:

The word 'cow' means ⟨cow-individuals, inherence relation, cowness⟩.

The first member of this meaning complex is the _śakya_, i.e., the referent or the referents of an expression, and the third member is the _śakyatāvacchedaka_, i.e., the limitor of the property of being the referent of an expression. In order to explain these terms we should draw distinctions among the terms '_śakya_' which stands for the referent or the referents of an expression, '_śakyatā_' which signifies the property of being the referent which resides in a referent, '_avacchedaka_' which means 'the limitor which determines a referent or a set of referents', '_avacchedakatā_' which signifies the property of being the limitor which resides in a limitor. In this context it is to be noted that by the word '_śakyatā_' or '_avacchedakatā_', the Nyāya specifies the role of an entity. In the above example the individual cows are considered as referents of the word 'cow'. Since they are playing the role of referents in this context, Nyāya has postulated an abstract property to characterize this role which is expressed by the word '_śakyatā_'. The property of being the referent which resides in the individual cow is limited by the property _cowness_ which is a limitor in this context. Since _cowness_ is playing the role of the limitor, it has the property of being the limitor. The relation of inherence which is the second member of this meaning-complex relates cowness to individual cows. Hence a cow-individual would be the first term (_anuyogin_), and the _cowness_ would be the second term (_pratiyogin_) of the relation of inherence.

Now the question is how the property of being the referent or the referents is related to the property of being the limitor. According to the Nyāya the former is _determined by_ (_nirūpita_) the latter and _vice versa_. Hence the property of being the referent (_śakyatā_) is _limited by_ a limitor, but _determined by_ (_nirūpita_) the property of being the limitor. In the above example the property of being the referent which resides in an individual cow is limited by the universal _cowness_, but determined by the property of being the limitor which resides in _cowness_.

The distinction between the relation _limited by_ and the relation _determined by_ may be explained in either of the following ways:

(I) _x_ is _limited by_ _y_ iff (a) _x_ and _y_ are properties, (b) _x_ is a relational property (_y_ may or may not be a relational property), and (c) the property _y_ is a mode of presentation of the object where the relational property _x_ resides.

(II) _x_ is _determined by_ _y_ iff both _x_ and _y_ are relational properties of correlatives.

In this context it is to be noted that the term 'limitor' has been used in various senses in the Nyāya literature. But in the context of meaning the term 'limitor' very closely corresponds to Frege's use of the term 'sense'. The limitor is a mode of presentation which uniquely determines the referent or the referents of a term. This use of the term 'limitor' is predominant in the later development of the Nyāya philosophy.

In this context I would like to mention that *śakyatāvacchedaka* which is the third member of the meaning-complex is considered as the reason or ground for applying a term to an object or objects to which it applies. In the technical language of the Nyāya this feature of the third member is expressed by the term '*pravṛtti-nimitta*'. The '*pravṛtti-nimitta*' or *śakyatāvacchedaka* of a term need not be a universal property or a class-character, but it is accepted as a universal in the case of a class-expression such as 'cow'. Since the reason for applying the term 'cow' to cow-individuals is the universal cowness, the *pravṛtti-nimitta* of the term 'cow' is considered as cowness. But the *pravṛtti-nimitta* of a term like 'animal' would be a complex property, i.e., the property of having a tail and a hairy body, which is not a universal or class-character (*jāti*). In the case of a definite description such as 'the father of John' the *pravṛtti-nimitta* would be a complex property, viz., the property of being the father of John. Similarly, in the case of a proper name such as 'John', according to some Nyāya philosophers, the *pravṛtti-nimitta* is the uniqueness of the individual John. As to the exact nature of *pravṛtti-nimitta* there is some difference of opinion among the Nyāya philosophers.

According to Vardhamāna Upādhyāya the definition of '*pravṛtti-nimitta*' of a term can be stated in the following way:

> *x* is the *pravṛtti-nimitta* of the term '*y*' iff (1) *x* is a part of the meaning-complex of the term '*y*', (2) *x* qualifies the referent or the referents of the term '*y*', and (3) *x* appears as the qualifier when we remember the referent or the referents of '*y*'.[57]

It is to be noted that this definition of the term '*pravṛtti-nimitta*' emphasizes both the epistemic and the semantic nature of it. At semantic level it is something which belongs to or qualifies the referent or the referents of a term, and at the level of thought it is a qualifier of a qualificand which corresponds to the referent or the referents of a term. To use the Fregean terminology, it 'belongs to the object' and it is 'a mode of presentation of the object.'

But this definition of Vardhamāna Upādhyāya has not been accepted by some followers of the Navya-Nyāya. The definition of Gadādhara, a Navya-Naiyāyika, is slightly different from that of Vardhamāna Upādhyāya. According to him the pravṛtti-nimitta of a term is a part of the meaning-complex, and it occurs in or belongs to the referent or the referents of a term which is another member of the meaning-complex.[58] Gadādhara's definition does not include the third condition of Vardhamāna's definition.

Raghunātha Śiromaṇi, a Navya Naiyāyika, has further simplified the definition of 'pravṛtti-nimitta'. He is willing to drop the first condition of pravṛtti-nimitta from Vardhamāna's definition of this term. That is to say, the pravṛtti-nimitta of a term need not be a part of the meaning-complex of it.[59] According to his view the pravṛtti-nimitta of a term is something which indicates the referent or the referents of a term. It is an indicator (upalakṣaṇa) of the object to which a term applies. The view of Raghunātha corresponds to the referential description of Kripke. As according to Kripke the term 'teacher of Alexander the Great' fixes the referent of the term 'Aristotle', so according to Raghunātha the term 'the property of being the locus of sound' fixes the referent of the term 'ākāśa' which is usually translated as 'sky' or 'ether'. Similar is the case with other terms such as 'cow'. Cowness which is the pravṛtti-nimitta of the term 'cow' is not part of the meaning-complex of it. Cowness indicates the referents of the term 'cow'. In other words, it also fixes the referents of the word 'cow'.

Our above discussion of pravṛtti-nimitta reveals certain striking similarities with the views of Frege and Kripke. The view of Vardhamāna closely corresponds with that of Frege, while the view of Raghunātha is similar to Kripke's method of fixing the referent of a name in terms of a referential definite description.

Now let us say a few words about the meaning of a sentence or a complex expression. Since the Nyāya explanation of the meaning of a sentence is analogous to that of the meaning of a complex term such as 'the author of Waverley' or 'a winged horse', this discussion would be useful for our purpose in the context of singular terms. In this context it is to be noted that the Nyāya has drawn distinctions among the terms 'sentence' ('vākya'), 'meaning of a sentence' ('vākyārtha'), and 'understanding the meaning of a sentence' ('vākyārthabodha' or śābdabodha', or 'anvayabodha').[60]

According to the Nyāya a sentence is an ordered set of meaningful expressions or morphemes which are related by syntactic rules. In other words, a sequence of morphemes having mutual syntactic expectancy (ākāṅkṣā) would constitute a sentence. The meaning of a sentence is not just the function

of the meanings of its parts which are expressions. Here we come across a striking similarity between Montague[61] semantics and the Nyāya semantics. The Nyāya claims that the meaning of a sentence, say Fa, is a function of the meanings of 'F', 'a', and the order of these two expressions which are related by syntactic rules. Moreover, the order of the morphemes will take care of the scope notations in terms of which we disambiguate an ambiguous expression.

Now let us consider the expression 'a blue pot' or 'a pot is blue'. The meaning of this expression would include the relation between the referents corresponding to the terms 'a pot' and 'blue'. This relation which is part of the meaning of every complex meaningful expression is due to the syntactic expectancy (ākāṅkṣābhāṣya)[62] between the expressions or morphemes which occur in a complex expression. In the above example it is a relation between a pot and a particular blue color, and this relation is called 'Inherence'. In this context it is to be noted that according to the Nyāya all relations are considered as dyadic. Hence a syntactic relation between the expressions would also be dyadic. All complex expressions would be considered as sets of ordered pairs, and all higher ordered sets would be reducible to a set of ordered pairs. Since all relations are dyadic and a complex expression is a set of ordered pair, the Nyaȳa can disambiguate an ambiguous expression of ordinary language, which is due to scope notations.

The Nyāya has drawn a distinction between 'the meaning of a sentence or a complex expression' and 'the understanding of the meaning of a sentence or complex expression'. The latter expression means 'the understanding of the relation which is due to the syntactic competency between the expressions' ('anvayabodha'). It refers to the cognition generated by a complex expression or a sentence. The conditions for understanding the meaning of a sentence would include semantic competency (yogyatā) between the referents of the expressions which occur in a sentence and the spatio-temporal contiguity or proximity (sannidhi) between these expressions in addition to the syntactic competency (ākāṅkṣā) between the expressions. For this reason, according to the Nyāya, the sentences like 'He irrigates the field with fire' do not generate any cognition in the hearer or the speaker, although they are meaningful sentences.[63] That is to say, we do not understand the meanings of such sentences, although they are meaningful sentences. From this distinction it follows that there is no one-to-one relation between the members of the sets corresponding to 'meaningful sentence' or 'sentence' and 'understanding the meaning of a sentence' (i.e., 'the sentences whose meaning we understand'). The set which corresponds to 'understanding of the meaning of a sentence' ('vākyārthabodha') would be a proper subset of the set which corresponds to

'sentence' ('vākya'). Moreover, according to the Nyāya, an identity sentence of the form 'a is a' or '(ιx) Fx is (ιx) Fx', and a contradictory sentence of the form 'An a which is F is not F' do not generate any cognition in a hearer or a speaker.[64] That is to say, we do not understand the meanings of such sentences. From this it does not follow that an identity or a self-contradictory sentence is meaningless. Both an identity sentence and a self-contradictory sentence are perfectly meaningful expressions, and each of them would receive a truth-value, i.e., truth or falsity.

Another important feature of the Nyāya theory of meaning of a sentence is that the relation which is part of the meaning of a sentence and which is due to the syntactic expectancy (ākāṅkṣābhāsya) between the expressions is not expressed or represented by any expression which occurs in that sentence. For example, the meaning-content of the sentence 'John is John' would include the relation of identity which is not mentioned by any expression. But the meaning-content of 'John is identical with John' would include an additional relation which would relate identical with John to John. In this case identity would be part of one of the terms of a self-linking (svarūpa) relation. On this point the Nyāya analysis is different from Frege's analysis of an identity sentence.[65] According to Frege the two sentences, viz., 'The morning star is Venus', and 'The morning star is no other than Venus' do not differ in meaning. But according to the Nyāya the meaning of the latter sentence is much richer than that of the former, although they have the same truth-value. Hence according to the Nyāya the meaning-content of a sentence is necessarily richer than the meanings of the terms which occur in a sentence or complex expression, and this richness is due to the mutual syntactic relation between the component parts of a sentence or complex expression. This element of meaning is the relation (anvaya) between the meanings of the terms of a sentence, and it is considered as the function of the syntactic expectancy (ākāṅkṣābhāsya) between the expressions of a sentence.

Our above discussion of meaning paves the way for the Nyāya discussion of proper names. In the Nyāya system alone we come across at least four different theories of proper names. Some of them would follow from our discussion of the meanings of terms like 'cow'. One of our aims in this paper is to substantiate one of the views of the Nyāya with some modification. Our aim is also to claim that this view is quite consistent with the sense theory of Frege and the baptism ceremony of Kripke.

(A) According to some of the early Nyāya philosophers including Jayanta[66] the meaning of a proper name is the object itself. Since a proper name is applicable to one object only, the meaning-complex cannot include a universal as it does in

the case of a general term such as 'cow'. The words like '<u>ā-kāsa</u>' ('sky' or 'ether'), and a proper name such as 'John' are included in this category. Hence the meaning of a proper name is an individual object which is devoid of any universal or class-character (<u>jāti</u>).[67] Here the word 'individual object' does not mean 'object along with its properties and relations', but rather 'object devoid of all its properties and relations'. So the meaning of a name such as 'Socrates' is the individual Socrates. The supporters of this view reject the thesis that the meaning of a proper name can be expanded in terms of a set of definite descriptions on the following grounds.[68]

If the meaning of the name 'Socrates' is equated with the meaning of 'the teacher of Plato', then the sentence 'Socrates is the teacher of Plato' becomes an identity sentence such that the mode of presentation or the limitor of the subject (<u>uddeśyatāvacchedaka</u>) is the same as the mode of presentation or the limitor of the predicate. In other words, the limitor of the referent which corresponds to the subject-expression would be the same as the limitor of the referent which corresponds to the predicate-expression. Hence the sentence 'Socrates is the teacher of Plato' would not be different from the sentence 'The teacher of Plato is the teacher of Plato'. According to the Nyāya if the limitor of the subject is the same as the limitor of the predicate, then the sentence does not generate any cognition. That is to say, neither the hearer nor the speaker of this sentence would understand the meaning of it as distinct from the meanings of its parts. This follows from the Nyāya theory of qualificative cognition.

According to the Nyāya cognitions are of two types, viz., qualificative and non-qualificative. In this context by the word 'cognition' we refer to the object or the content of a cognition. A qualificative cognition can be expressed by a complex expression of the form '<u>aRb</u>'. A qualificative cognition involves necessarily at least three <u>epistemic</u> elements, viz., a qualificand, a qualifier, and a relation between them. The relation between the qualificand and the qualifier at cognitive level is called the 'qualification relation' ('<u>viś-esya-viśesana-sambandha</u>'). At the ontological level there might not be three elements corresponding to a qualificative cognition. Let us consider the cognition of a pot, which is one of the simplest cognitions. The content of this cognition has a pot as its qualificand, potness as its qualifier, and the relation of inherence as its qualification relation. Since a pot is cognized under a mode of presentation, this cognition is considered as a qualificative cognition. Since a relation relates the qualifier to the qualificand, this type of cognition is also called 'relational cognition'. Even the cognition represented by a proper name,

say John, would contain three elements at the epistemic level. These three elements would be the individual John, the property of being John (tadvyaktitva), and the self-linking (svarūpa) relation. Since a self-linking relation is not a separate ontological entity, it has to be identified either with the first term (anuyogin) or with the second term (pratiyogin). However, it is usually identified with the first term. So at the ontological level we do not have three distinct entities. But in a non-qualificative cognition the ultimate elements of a qualificative cognition are cognized without any mode of presentation. If 'aRb' represents the cognition of a pot, then in a non-qualificative cognition the referents of 'a', 'b', and 'R' are cognized without any mode of presentation.

Now let us consider the cognition represented by the sentence 'a pot is blue'. Here a pot is the qualificand, a blue color is the qualifier, and the relation of a blue color to a pot is the relation of inherence. If the qualificand and the qualifier are different, then they must have different limitors. In this case potness is the limitor of the property of being the qualificand which resides in a pot, and blueness is the limitor of the property of being the qualifier which resides in a blue color. Here both a pot and a blue color are contents of a cognition. The relation of inherence which relates a blue color to a pot is called 'the limiting relation of the property of being the qualifier' ('prakāratā-avacchedaka-sambandha'). So the property of being the qualifier which resides in a qualifier is limited by both a property and a relation, but the property of being the qualificand which resides in a qualificand is limited by a property alone. According to the Navya-Nyāya if the qualificand and the qualifier are different, then their property-limitors must also be different. Let us consider the sentence 'a pot is a pot'. If this sentence were to represent a cognition, then the qualificand would also be the same as the qualifier, and the mode of presentation of the qualificand would also be the same as the mode of presentation of the qualifier. That is to say, the property (dharma) which would limit the property of being the qualificand would be the same as the property (dharma) which would limit the property of being the qualifier. So the property-limitor (avacchedaka-dharma) would be the same in both the cases. Hence the sentence 'a pot is a pot' would not represent a cognition different from the cognition expressed by the term 'a pot'. For this reason there cannot be understanding of the meaning (śābdabodha) of the sentence 'a pot is a pot', although there is understanding of the meaning of the term 'a pot'. Similarly, the sentence 'The teacher of Plato is the teacher of Plato' would not generate a cognition as distinct from the cognition generated by the expression 'the teacher of Plato.' In other words, the contents of these two

cognitions would not be different. For this reason the Nyāya claimed that there is no understanding of the meaning of the sentence 'The teacher of Plato is the teacher of Plato', although there is understanding of the meaning (śābdabodha) of the term 'the teacher of Plato'. Similar is the case with the sentence 'John is John'. There is no understanding of the meaning of this sentence as distinct from the understanding of the meaning of its parts. Hence according to the Nyāya an identity sentence of the form 'a is a' or '(ιx) Fx is (ι) Fx', or an identity sentence with the same indefinite description as its terms such as 'a pot is a pot' does not generate a cognition as distinct from the cognition or the cognitions generated by its parts. Each part of an identity sentence would generate the same type of cognition. Hence their contents would be the same. From this view of the Nyāya one should not conclude that an identity sentence of the above form is meaningless, neither true nor false. According to the Nyāya each sentence of the above form would be perfectly meaningful, and it would be either true or false.

From the above discussion of identity sentences it follows that the content of the cognition generated by the sentence 'Socrates is the teacher of Plato' is not different from that of the cognition generated by the term 'the teacher of Plato', if the meaning of 'Socrates' is 'the teacher of Plato'. This situation can be avoided if it is claimed that the meaning of a proper name is the object itself. That is to say, the meaning-complex of a proper name includes only the referent of it. On this point the view of Jayanta and his followers is similar to Russell's view of logically proper names. But the followers of the Navya-Nyāya have raised certain objections against this view. Some of the objections are similar to those raised against Russell's conception of a logically proper name. Let us discuss some of these objections.

(1) If the meaning of a proper name such as 'Tom' is the referent itself, then sentences like 'Tom is honest' or descriptive expressions such as 'a brother of Tom' would not generate any cognition in a hearer or a speaker if he is not acquainted with the referent of 'Tom'. In other words, a speaker or a hearer would not understand the meaning of an expression of this type. Even if the speaker is acquainted with the referent of 'Tom', the hearer would not understand the meaning of 'Tom is honest' or 'a brother of Tom'. Hence such sentences or expressions cannot be used to communicate the thoughts or the cognitions of the speaker. In defense of this view of the old Nyāya it may be said that the speaker himself can write down such sentences or expressions for his personal use. He might keep a diary of his thoughts by writing down sentences or expressions of this form. Against this defense it may be said that the speaker also would not understand the meanings of such expressions if the referent of a

proper name is destroyed or if he is not acquainted with the referent. From this observation the followers of the Navya-Nyāya conclude that this type of proper name is not useful either to a speaker or to a hearer.

(2) Secondly, this view of Jayanta and his followers goes against the Nyāya thesis that only a class-character (jāti) and an unanalyzable imposed property (akhanda-upādhi) can be cognized without any mode of presentation (i.e., without any property-limitor (avacchedaka-dharma)) unless they have been mentioned by expressions.[69]

In order to understand this remark let us explain the Nyāya conception of a property and the Nyāya analysis of sentences such as (1) John is a man, and (2) John has humanity.

According to the Nyāya anything which has a locus is considered as a property of that locus. So a property may be defined in the following way:

x is a property = df $(\exists y)$ (y is a locus of x).[70]

Since a class-character (jāti), quality (guna), and action (karma) inhere in a substance, they are considered as properties of substances.[71] The particularities of eternal substances are considered as properties of those substances. Since a whole (avayavin), such as a table, inheres in its parts (avayava), a whole is considered as a property of its parts. Similarly, the property of being John (i.e., the uniqueness of John) is considered as a property of John. This property is different from a class-character, quality, action, and particularity of an eternal substance (antya-viśesa). Since the property of being John has no parts, it is called an 'unanalyzable imposed property'. By using the lambda operator it may be symbolized in the following way:

(λx) (John=x).[72]

But the property of being an animal (paśutva) is an analyzable imposed property (sakhanda-upādhi), and it is reducible to the property of having a hairy body and a tail. Similarly, the property of being the eldest son of John is considered as an analyzable imposed property. It may be symbolized in the following way:

(λy) $(((\iota x)\ Fx) = y)$,
where '$(\iota x)\ Fx$' stands for 'the eldest son of John'.

Now let us consider the difference between the sentences 'John is a man' and 'John has humanity'. In the cognitive counterpart of the former sentence John is the qualificand and

a man is the qualifier. The qualificand, according to the Nyāya, must be presented under some mode of presentation (avacchedaka), and so must be the qualifier. In this case the qualificand is presented under the mode of the property of being John, and the qualifier is presented under the mode of the class-character humanity. Since these modes of presentations or limitors are unanalyzable properties and not mentioned by any expression, they do not require any further mode of presentation. But in the cognitive counterpart of the sentence 'John has humanity', John is the qualificand, and the class-character humanity is the qualifier. Since a class-character is mentioned by an expression, it must have a mode of presentation (avacchedaka). The mode of presentation of the qualifier humanity is the property of being humanity which is nothing but the property of being present in all and only human beings by the relation of inherence. So the class-character humanity is presented under the mode of the property of being present in particular human beings alone, which is a very complex property.[73] Moreover, the qualification relation would not be the same in both the cases. In the former case it would be the relation of identity, while in the latter case it would be the relation of inherence. In this way the Nyāya has drawn the distinction between the contents of these two sentences.

Now if we follow Jayanta's view of proper names, then we have to admit that an individual such as John can be presented without any mode of presentation even if he is mentioned by an expression. So we are going against the thesis that if anything is mentioned by an expression, then the cognition of it requires a mode of presentation. Moreover, this view would multiply the number of entities which can be cognized without any mode of presentation. The list would include the referents of proper names in addition to the class-characters and the unanalyzable imposed properties which have not been mentioned by any expression. For these reasons the followers of the Navya-Nyāya have rejected this theory of proper names.

(B) Now let us consider the views of the followers of the Navya-Nyāya. (a) Some of the followers of the Navya-Nyāya such as Jagadīśa[74] have claimed that there is no fundamental difference between a proper name and a general term such as 'man'. As the term 'man' refers to a set of particular human beings under the mode of humanity, so the name 'John' refers to a set of John-individuals under the mode of the property of being John. According to this view the person called by the name 'John' is a series of John-individuals, and the term 'John' applies to each of the members of this series. Hence a proper name such as 'John' is on a par with a general term such as 'man'. In this respect the Nyāya use of a proper name is similar to Russell's use of the term 'Piccadilly'.[75]

Moreover, according to this view as the meaning-complex of a general term, say 'cow', contains three elements, viz., (1) the referents of the word 'cow', (2) a unique and an essential property of the referents, which is cowness in this case, and (3) the relation of inherence which relates the latter to the former, so the meaning-complex of a proper name such as 'John' contains three elements, viz., (1) the referents of the word 'John', i.e. different John-individuals, (2) a unique and an essential property of the referents, which is Johnhood in this case, and (3) the relation of inherence which relates the latter to the former. In this context the word 'John-individuals' refers to different stages of John such as the John of boyhood, the John of adulthood or the John of old age.

The difference between the word 'cow' and the word 'John' has been emphasized in terms of baptism ceremony. According to Jagadīśa in the case of a proper name such as 'John' the baptism or name-giving ceremony takes place at a particular time, and the father of John says, "Let the word 'John' generate the cognition of my son."[76] So the relation between the word 'John' and the meaning-complex is set up by the intention or the desire of a particular person, and the subsequent users of this name should use the word 'John' to refer to the same John-individuals. If the relation remains the same, then the referents of 'John' would also remain the same.

But in the case of the word 'cow' the relation has been set up by the intention of God. This is another way of saying that the origin of this baptism ceremony is not known. but in both cases the meaning-relation is set up by the intention of the name-giver. So the Nyāya emphasises the conventional relation between an expression and what is referred to by that expression. In this respect we find a striking similarity between the view of the Nyāya and that of Kripke. Both of them emphasize the initial baptism and the intention of the name-giver.

Certain objections can be raised against this view of the Navya-Nyāya.

(1) First of all, this view ignores the distinction between a general term and a proper name. Hence it fails to explain the difference between the relation which holds among the referents of a general term such as 'cow', on the one hand, and the relation which holds among the referents of a proper name such as 'John' on the other. Different referents of the word 'cow' are not dependent on each other in the way the referents of the word 'John' are dependent on each other. The referents of the word 'John' are different John-individuals such as the John of boyhood, the John of adulthood, and the John of old age. The John of adulthood is dependent on the John of boyhood, and the John of old age is dependent on the John of adulthood. But a cow in Hawaii is not dependent

on a cow in California in the same way. Hence the relation between the boy-John and the adult-John is not the same as the relation between two particular cows. In the former case the concept of sequence is present, but not in the latter case. This distinction cannot be explained if a proper name is assimilated to a general term.

(2) Secondly, this view of the Navya-Nyāya cannot avoid the problem of proper names by treating ordinary proper names as general terms. It simply pushes the problem one step further. Since according to the Nyāya everything is nameable,[77] we can name each of the members of the John-individual series. Instead of giving one name to a series we can give a separate name to each member of the series. If these names are also treated as general terms, then this view will lead to an infinite regress, and it becomes indefensible.

(3) Thirdly, one of the consequences of this view is that we have to accept one class-character (jāti) corresponding to each proper name. This will go against the law of parsimony which is so widely used in the Nyāya literature. Moreover, we have to admit the being of these class-characters even before the referents of proper names come into being. For these reasons some of the followers of the Navya-Nyāya do not consider proper names as general terms.

(b) The second view of the Navya-Nyāya[78] does not treat proper names as general terms, rather it considers proper names as names for particular objects or persons. According to this view the meaning-complex of a proper name such as 'John' would include the individual John, the property of being John (tadvyaktitva), and the relation of the latter to the former. The property of being John would be the mode of presentation of John, and hence it would be the limitor of the property of being the referent (śakyatā) which resides in John. the meaning-complex may be represented in the following way:

$$\text{'John' means } \langle \text{John}, R, (\lambda x)(\text{John}=x) \rangle .$$

The third member of this ordered triple is the qualifier (viśesana) which qualifies the first member which is the qualificand (viśesya), and the R is the relation of the qualifier to the qualificand, which is a self-linking (svarūpa) relation in this context. A self-linking relation is not an ontological entity or category. If one term is related to another by a self-linking relation, then it is to be identified with one of the terms. It is usually identified with the first term.

As regards the nature of the property of being John it is said that it cannot be identified with any property expressible by a definite description such as 'the son of Tom'. Moreover, it cannot be identified with a conjunction or

disjunction of properties of John. It is considered as a necessary and unique property of John. Now the question is what this property could be. Since it is not considered as something distinct from all the seven categories of the Nyāya-Vaiśeṣika system, it is to be identified with some category or the other. Jagadīśa claimed that the property of being John is the individual John itself as the second term of the relation of identity whose first term is also John. According to the Nyāya every object has its own relation of identity. There are as many identity relations as there are objects in the Nyāya system. It is not the same relation which relates an object to itself. The identity relation of John to himself is different from that of Tom to himself.[79] If a person such as John is related to himself by the relation of identity, then John is both the qualificand and the qualifier of this relation. The property of being the qualifier resident in John is limited by the relation of identity (tādātmyasambandhāvacchinna John).[80] So virtually this view of Jagadīśa claims that the same individual could be conceived both as a qualificand and qualifier in such cases.

Even if this thesis of Jagadīśa and his followers can be considered as a limiting case of the qualificand-qualifier relation, it does not solve the problem of communicability of the meaning of a proper name. How can a hearer who is not acquainted with John understand the meaning of the word 'John'? Since the sense of the word 'John' is explained in terms of John, a speaker or a hearer would not undertand its meaning unless he is acquainted with John. Moreover, since the property of being John which is the mode of presentation of John is considered as the individual John itself as the second term of the relation of the identity, it goes against the Nyāya thesis that there is a difference between an object and the mode in which it is presented to us. Moreover, if we admit that an object including a class-character can be its own mode of presentation, then there is no need to admit any higher-order abstract entities such as cownessness, i.e., the property of being cowness (gotvatva) which is the mode of presentation of cowness. Since the Nyāya has explained the property of being cowness as the property of being present in all and only cows by the relation of inherence,[81] it cannot be explained in the way the property of being John has been explained. If it is claimed that the property of being cowness can be explained in either of the two ways, then one of them would be superfluous unless we point out some specific reason for both of them.

(c) The third view of proper names is associated with Raghunātha Śiromaṇi,[82] a follower of the Navya-Nyāya. His view differs from the second view in two important respects. Raghunātha claimed that the pravṛtti-nimitta, i.e., the limitor of the property of being the referent of a term is not

part of the meaning-complex of a term.[83] Hence the meaning of a term such as 'John' would be the referent, and the mode of presentation of the referent (śakyatāvacchedaka) is something which indicates or suggests (upalakṣaṇa) the referent. Raghunātha's conception of the mode of presentation of the referent corresponds to Kripke's conception of referential description. It is a device for fixing the the referent or the referents of a term. In the case of a proper name such as 'John' the mode of presentation of John is the property of being John (tadvyaktitva). This property is something which fixes the referent of 'John', and it does not form part of the meaning-complex of the name 'John'. As regards the nature of the property of being John, Raghunātha claimed that it cannot be identified either with John or with any other category of the Nyāya-Vaiśeṣka system.[84] The property of being John is a separate irreducible category by itself.

The merit of this view is that it does not identify the property of being John with John. In spite of this merit this view is not acceptable to most of the followers of the Navya-Nyāya. It is claimed that this view violates the law of parsimony, and populates the Nyāya ontology with a number of irreducible entities. So we are landed with two extreme views. One of them equates the property of being John with John-individual as the second term of the relation of identity, and the other one considers it as an irreducible entity. In the midst of this controversy I would like to suggest a mean between the two extremes without going against any fundamental thesis of the Nyāya. It seems to me that this view can be developed within the Nyāya system.

(d) According to this view the property of being John which is the reason or the ground for applying the term 'John' to the individual John is not a spearate category or entity. It is a function of the uniquenesses of the parts of John. The parts of an individual can also be named, and the uniqueness of each of the parts is to be explained in terms of the uniquenesses of its parts. But this process cannot go on *ad infinitum* in the Nyāya system. According to the Nyāya any created object such as a table, a city, or the body of a person is ultimately reducible to a set of ultimate atoms, each of which is characterized by an ultimate differentia (antya viśeṣa). But according to Raghunātha this process of reduction should terminate at a tryaṇuka which is the minutest visible particle consisting of three dyadic atoms.[85]

The differentia of an ultimate substance, according to the Nyāya, is a separate ontological category. An ultimate differentia differentiates an ultimate substance, and it differentiates itself from all other objects. So the differentia of an atom distinguishes that atom from all other atoms, and it distinguishes itself from everything else. The objects which consist of atoms or ultimate entities are determined by

the uniqueness of their parts. That is to say, the uniquenesses of a whole can be explained in terms of the uniquenesses of its parts. Since the Nyāya has admitted ultimate differentiae of the irreducible parts, there is no need to assign a separate uniqueness to each of the constructed or created objects. The uniqueness of the whole is a function of the uniquenesses of its parts. Now the question is how the term 'function' is to be understood in this context. It might be suggested that as in Frege the sense determines the referent of a term, so in this case the uniquenesses of the parts determine the whole or the referent of a term. For example, the city of Honolulu is determined by the uniqunesses of its parts which are nothing but different areas of the city. These parts are, in turn, determined by the uniquenesses of their parts, and this process goes on until we arrive at the irreducible parts of a system. If the irreducible parts are visible objects, then each minutest visible object has a uniqueness, and the objects which are constructed out of such irreducible parts are determined by their uniquenesses. Another way of describing this situation is to consider a function as a many-one relation. This is also an explication of the concept of determination. The uniquenesses of the parts are in the domain of this many-one relation, and the objects such as Honolulu or the body of a person such as John is in the counterdomain of this relation. In this context it is to be noted that our domain-counterdomain distinction is relative to a context of discourse. What follows from this view is that the uniqueness of a whole is explained in terms of its structural properties, i.e., the uniquenesses of its parts. Moreover, this view gives us a criterion for identifying the referent of a proper name in terms of the uniquenesses or properties of its parts without postulating a separate irreducible uniqueness for the whole. It seems to me that this explanation of the uniqueness of an object is in accord with Kripke's rigid definite descriptions which are given in terms of the structural properties of objects. So in a particular discourse the referent of a name, say 'John', is to be identified in terms of some uniqueness or uniquenesses of the parts of John's body. But if we want to identify John in all possible worlds, then the uniquenesses of the irreducible parts of John would be considered as both necessary and sufficient conditions. Now the question is whether the property of being John which is explained in terms of the uniquenesses of his parts is to be included within the meaning-complex of the term 'John'. On this point we can follow either the main stream of the Nyāya philosophy or the view of Raghunātha. If we follow the latter, then the meaning-complex of 'John' would not include the property of being John. It would be considered as something which fixes or indicates (upalakṣaṇa) the referent of the word 'John'. But if we follow

the main stream of the Nyāya philosophy, then, unlike Kripke, the Nyāya would include the property of being John which is explained in terms of the uniquenesses of the parts of John within the meaning-complex of the word 'John'. This would give us the reason for applying a proper name to the object to which it applies. In spite of this difference both Kripke and the Nyāya would agree on some other points. Both of them have emphasized the baptism ceremony and the intention of the name-giver which sets up the relation between a proper name and the object referred to by it. But the Nyāya would insist on the method of identifying the referent of a name. This epistemic conditon, according to the Nyāya, is to be included in the meaning-complex of a name. This epistemic element is the Fregean element of sense. Hence the Nyāya view contains both the Kripkean and Fregean elements so far as the meaning of a proper name is concerned. Instead of being incompatible with both, the Fregean and the Kripkean elements of meaning have been combined in the Nyāya concept of the meaning-complex of a proper name.

NOTES

1. Dummett, M., Frege: Philosophy of Language, p. 54.
2. Ibid., p. 55.
3. Frege, G., Translations from the Philosophical Writings of Gottlob Frege, edited by P. Geach and Max Black, p. 43.
4. Ibid., p. 45.
5. Ibid., p. 45.
6. Ibid., p. 49.
7. Ibid., p. 49.
8. Frege, G., 'On Sense and Nominatum', Readings in Philosophical Analysis, edited by H. Feigl and W. Sellars, pp. 85-86.
9. Ibid., p. 85.
10. Ibid., p. 87.
11. Dummett, M., Frege, p. 90.

12. Frege, G., 'On Sense and Nominatum', Readings in Philosophical Analysis, edited by Feigl and Sellars, p. 93.

13. Ibid., pp. 85-86.

14. Ibid., p. 86.

15. Ibid., p. 86.

16. Ibid., p. 86.

17. Ibid., p. 86.

18. Frege, G., Posthumous Writings, translated by P. Long and R. White, p. 125.

19. Russell, b., Introduction to Mathematical Philosophy, p. 173.

20. Ibid., p. 174.

21. Russell, B., 'The Philosophy of Logical Atomism', reprinted in Readings in Twentieth-Century Philosophy, edited by W.P. Alston and G. Nakhinikian, p. 358.

22. Ibid., p. 306.

23. Ibid., p. 303.

24. Ibid., pp. 308-309.

25. Ibid., p. 309.

26. Ibid., p. 316.

27. Ibid., p. 320.

28. Kripke, S., 'Naming and Necessity', in Semantics of Natural Language, edited by D. Davidson and G. Harman, p. 257.

29. Ibid., p. 257.

30. Ibid., p. 258.

31. Ibid., p. 285.

32. Ibid., p. 294.

33. Ibid., p. 296.

34. Ibid., p. 296.

35. For a criticism of Kripke on this point, see my paper 'Some Reflections on Kripke', *Logique et Analyse*, 1980.

36. Kripke, S., 'Naming and Necessity', in *Semantics of Natural Language*, p. 280.

37. Ibid., pp. 298-299.

38. Ibid., p. 300.

39. Ibid., p. 301.

40. Ibid., p. 301.

41. Ibid., p. 302.

42. Frege, G., 'On Sense and Nominatum', in *Readings in Philosophical Analysis*, p. 86.

43. Ibid., p. 86, footnote 2.

44. Ibid., p. 86.

45. M. Dummett in his book *Frege*, and G. Evans in his article 'Causal Theory of Names' have raised several objections against Kripke's theory of names.

46. Some of the remarks mentioned in this context are taken from my paper 'Some Reflections on Kripke', *Logique et Analyse*, 1980.

47. *asmāc chavdād ayam artho boddhavyah*, quoted in *Śabdaśaktiprakāśikā* of Jagadīśa, Bengali translation and commentary by Pandit Madhusudana Nyāyāchārya, *Our Heritage*, 1974, p. 99.

48. *idam padam imam artham bodhayatu*, ibid., p. 99.

49. In this context I have followed the classification of Pandit Madhusudana Nyāyāchārya. But Jadadīśa has given a slightly different characterization of '*paribhāsā*', compare, *Our Heritage*, 1974, p. 99, with *Our Heritage*, 1976, pp. 30-33.

50. In this context it is to be noted that our use of the term 'referent' simply means 'the second member of the meaning-relation'. In the sequel we shall discuss the nature of this second member.

51. jātyākṛtivyaktayaḥ padārthaḥ, Nyāyasūtra, quoted in Jagadīśa, Our Heritage, 1977, p. 48.

52. Since the singular suffix is used in the Nyāyasūtra, the meaning-relation is taken as one.

53. I am indebted to Pandit Visvabandhu Tarkatirtha for this interpretation of the word 'ākṛti'.

54. Jagadīśa's Śabdaśaktiprakāśikā, Our Heritage, 1977, p. 51.

55. Ibid., p. 50.

56. Ibid., p. 58.

57. vācyate sati vācyavṛttitve sati vācyopasthitiprakāratvam, Our Heritage, 1976, p. 35.

58. vācyate sati vācyavṛttitva, Gadādhara's Śaktivāda.

59. According to Raghunātha the term 'vācyatve sati' should not be included in the definition of 'pravṛtti-nimitta', Our Heritage, 1976, p. 35.

60. Pandit Madhusudana Nyāyāchārya, 'Śabda-Prāmānya', Our Heritage, 1970, pp. 49-70. See also Jagadīśa's Śabdaśaktiprakāśikā, and Jogendranath Bagchi's Vākyārtha nirūpaner dārśanika paddhati.

61. R.H. Thomason (ed.), Formal Philosophy: Selected Papers of Richard Montague.

62. See Gadādhara's Śaktivāda.

63. See Jagadīśa's Śabdaśaktiprakāśikā, Our Heritage, 1975, p. 145.

64. Here we are not discussing āhāryajñāna.

65. Frege, G., Translations from the Philosophical Writings of Gottlob Frege, p. 44.

66. Jayanta, Nyāyamañjarī, edited by Sūryanārāyana Śukla, p. 298.

67. abhidheyasya sāmānyaśūnyatvād vyaktivācitā, ibid., p. 298.

68. In this context I have utilized some of the materials from Pandit Visvabandhu Tarkatīrtha's paper on 'Proper Names', which is written in Bengali.

69. Nyāyaratna, Maheśa Chandra, Navya-Nyāya-Bhāsāpradīpah, edited with commentary and Bengali translation by Kalipada Tarkāchārya, pp. 4-6.

70. Ibid., pp. 1-2.

71. Here the terms 'substance', 'quality' and 'action' are to be understood in the light of the Nyāya definition of these terms.

72. McCawley, J.D., Everything that Linguists Have Always Wanted to Know About Logic, p. 395.

73. For a detailed discussion of this property, see J.L. Shaw, 'The Nyāya on Cognition and Negation', Journal of Indian Philosophy, 1980.

74. Śabdaśaktiprakāśikā, Our Heritage, 1975, p. 178.

75. Russell, B., 'The Philosophy of Logical Atomism', reprinted in Readings in Twentieth Century Philosophy, edited by P. Alston and G. Nakhnikian, pp. 308-309.

76. Jagadīśa has used the word 'Caitra'. Cf. Caitrapadam matputragocarabodhajanakatāvat bhavatu, Our Heritage, 1976, p. 236.

77. For a comprehensive discussion on 'Nameability' see J.L. Shaw, 'The Nyāya on Existence, Knowability and Nameability', Journal of Indian Philosophy, 1978.

78. This view has been mentioned by Pandit Visvabandhu Tarkatīrtha, and it can be reconstructed from Jagadīśa's discussion of the name 'Dittha' which is a name for a wooden elephant.

79. Śabdaśaktiprakāśikā of Jagadīśa, Our Heritage, 1972, p. 16.

80. tādātmyena saiva vyaktis tadvyaktitvam, Paksatā of Jagadīśa, quoted in The Padārthatattva-Nirūpanam of Raghunātha Śiromani, edited by Pandit Madhusudana Nyāyachārya, p. 39. abhedastādātmyam, tacca svavrttyāsādharano dharmah tādrśo dharmas tat-tadvyaktitvādirūpa eveti, Vyutpattivādah of Gadādhara, quoted in the same book, p. 39.

81. See J.L. Shaw, 'The Nyāya on Cognition and Negation', Journal of Indian Philosophy, 1980, pp. 280-282.

82. Raghunātha Śiromani, The Padārthatattva-Nirūpanam, edited with Bengali translation and elaborate exposition by Pandit Madhusudana Nyāyāchārya, p. 39.

83. See commentary on Jagadīśa's Śabdaśaktiprakāśikā, Our Heritage, 1976, p. 35.

84. See Pandit Madhusudana Nyāyāchārya's commentary on The Padārthatattva-Nirūpanam, p. 39.

85. Ibid., p. 16 and p. 70.

REFERENCES

Alston, W.P. and G. Nakhnikian (eds.): 1963, Readings in Twentieth Century Philsophy, Free Press, New York.

Angelelli, I.: 1967, Studies on Gottlob Frege and Traditional Philosophy, D. Reidel, Dordredcht.

Bagchi, J.: 1981, Vākyārtha Nirūpaner Dārśanika Paddhati, Sanskrit Book Dipo, Calcutta.

Bell, D.: 1979, Frege's Theory of Judgment, Clarendon Press, Oxford.

Bhattacharya, P. (ed. and tr.): Bengali year 1377, Bhāṣāparicchedah with Siddhānta-muktāvalī, Sanskrit Pustaka Bhandar, Calcutta.

Bhattacharya, S.: 1976, 'Some Principles and Concepts of Navya-Nyāya Logic and Ontology', in Our Heritage, Sanskrit College, Calcutta.

Davidson, D. and G. Harman (eds.): 1972, Semantics of Natural Language, D. Reidel, Dordrecht.

Devitt, M.: 1981, Designation, Columbia University Press, New York.

Dummett, M.: 1973, Frege Philosophy of Language, Duckworth, London.

Dummett, M.: 1978, Truth and Other Enigmas, Harvard University Press, Cambridge, Mass.

Evans, G.: 1982, The Varieties of Reference, (ed.) John McDowell, Oxford University Press, Oxford.

Feigl, H. and W. Sellars (eds.): 1949, Readings in Philosophical Analysis, Appleton-Century-Crofts, New York.

Frege, G.: 1979, Posthumous Writings, tr. by Peter Long and Roger White, Basil Blackwell, Oxford.

Frege, G.: 1980, Translations from the Philosophical Writings of Gottlob Frege, 3rd edition, (eds.) P. Geach and M. Black, Basil Blackwell, Oxford.

French, P.A., Uehling, T.E. and H.K. Wettstein (eds.): 1979, Contemporary Perspectives in the Philosophy of Language, University of Minnesota Press, Minneapolis.

Gadādhara: 1929, Śaktivāda, (ed.) G.D. Sastri, Kashi Sanskrit Series, Chowkhamba, Varanasi.

Geach, P.T.: 1962, Reference and Generality, Cornell University Press, Ithaca, New York.

Geach, P.T.: 1972, Logic Matters, Basil Blackwell, Oxford.

Grossmann, R.: 1973, Ontological Reduction, Indiana University Press, Bloomington.

Ingalls, D.H.H.: 1951, Materials for the Study of Navya-Nyāya Logic, Harvard University Press, Cambridge, Mass.

Jagadīśa: 1972-1977, Śabdaśaktiprakāśikā, parts I and II, (ed. with Bengali translation and commentary by Madhusudan Nayāyacharya, Sanskrit College (Our Heritage), Calcutta.

Jayanta: 1971, Nyāyamañjarī, (ed.) Pt. Surya Narayana Sukla, Chowkhamba Sanskrit Series, Varanasi.

Klemke, E.D. (ed.): 1968, Essays on Frege, University of Illinois Press, Chicago.

Klemke, E.D.: 1971, Essays on Bertrand Russell, University of Illinois Press, Chicago.

Kunjunni Raja, K.: 1963, Indian Theories of Meaning, Adyar Library Series, Madras.

Matilal, B.K.: 1968, The Navya-Nyāya Doctrine of Negation, Harvard University Press, Cambridge, Mass.

Matilal, B.K.: 1971, Epistemology, Logic and Grammar in Indian Philosophical Analysis, Mouton, The Hague.

McCowley, J.D.: 1981, Everything That Linguists Have Always Wanted to Know About Logic, University of Chicago Press, Chicago.

Montague, R.: 1974, Formal Philosophy, (ed.) R.H. Thomason, Yale University Press, New Haven.

Nyāyaratna, M.C.: 1973, Navya-Nyāya-Bhāṣāpradīpah, (ed.) K. Tarkāchārya, with the commentary Suprabhā and Bengali translation, Sanskrit College, Calcutta.

Platts, M.: 1979, Ways of Meaning, Routledge and Kegan Paul, London.

Platts, M. (ed.): 1980, Reference, Truth and Reality, Routledge and Kegan Paul, London.

Prior, A.N.: 1971, Objects of Thought, (eds.) P.T. Geach and A.J.P. Kenny, Clarendon Press, Oxford.

Prior, A.N.: 1976, The Doctrine of Propositions and Terms, (eds.) P.T. Geach and A.J.P. Kenny, Univerity of Massachusetts Press, Amherst.

Quine, W.V.O.: 1972, Methods of Logic, Holt, Rinehart and Winston, New York.

Quine, W.V.O.: 1960, Word and Object, MIT Press, Cambridge, Mass.

Raghunātha Śiromaṇi: 1976, The Padārthatattva-Nirūpanam of Raghunātha Śiromani, with Bengali translation and elaborate exposition, ed. Pandit Madhusudana Nyāyācharya, Sanskrit College, Calcutta.

Russell, B.: 1940, An Inquiry into Meaning and Truth, George Allen and Unwin, London.

Russell, B.: 1948, Human Knowledge: Its Scope and Limits, Simon and Schuster, New York.

Russell, B.: 1919, Introduction to Mathematical Philosophy, Simon and Schuster, New York.

Schwartz, S.P. (ed.): 1977, Naming, Necessity and Natural Kinds, Cornell University Press, Ithaca.

Searle, J.: 1958, 'Proper Names', Mind, reprinted in Philosophy and Ordinary Language (ed.) C.E. Caton, Urbana, Illinois, 1963.

Shaw, J.L.: 1974, 'Empty Terms: The Nyāya and the Buddhists', Journal of Indian Philosophy 2, pp. 255-266.

Shaw, J.L.: 1978, 'The Nyāya on Existence, Knowability and Nameability', Journal of Indian Philosophy 5, pp. 255-266.

Shaw, J.L.: 1980, 'The Nyāya on Cognition and Negation', Journal of Indian Philosophy 8, pp. 279-302.

Shaw, J.L.: 1980, 'Some Reflections on Kripke', Logique et Analyse 23, pp. 345-363.

Sluga, H.D.: 1980, Gottlob Frege, Routledge and Kegan Paul, London.

Strawson, P.F.: 1959, Individuals, Methuen, London.

Strawson, P.F.: 1974, Subject and Predicate in Logic and Grammar, Methuen, London.

Tarkatirtha, V.: (unpublished), 'Proper Names', written in Bengali.

Thiel, C.: 1965, Sense and Reference in Frege's Logic, D. Reidel, Dordrecht.

Wittgenstein, L.: 1963, Tractus Logico-Philosophicus, tr. D.F. Pears and B.F. McGuinness, Routledge and Kegan Paul, London.

Wittgenstein, L.: 1958, Philosophical Investigations, Macmillan, New York.

Bimal Krishna Matilal

AWARENESS AND MEANING IN NAVYA-NYĀYA

1. AWARENESS OF THE OBJECT AS QUALIFIED

In this paper, I shall try to formulate and examine two interrelated Nyāya principles about knowledge, and show how a tension evolved in the system as a result of their interaction. These two fundamental principles had far-reaching effects in shaping the Nyāya theory of meaning as well as its theory of verbal knowledge derived from linguistic utterance. This latter part, however, will not be fully discussed here, for I shall devote this paper mainly to the clarification of the said principles and what, if anything, could be done with them.

The first principle that I shall take up for examination could be very roughly stated as follows: Whenever an object x figures (or floats, or swims = avagāhate) in my awareness, it figures there as something, i.e., under some guise or mode, distinguished in some way or other. The second principle may also be stated, using similar metaphorical expressions as is usual with the Sanskrit writers, as follows: A pre-condition for having a clear and distinct, i.e. qualificative, awareness of the above kind is a further awareness of the qualifier or distinguishing element (viśesana). To indulge in the elaboration of the implicit metaphor, we may say: When an object x features or floats distinctly in our awareness, it is invariably distinguished by a 'cloak' that may either be put upon it by us, or that may belong there initially, and be recognized by us as such. The second principle demands that to have such a distinguishing awareness (and a piece of knowledge is only a special case of such a distinguishing or qualificative or 'propositional' awareness in this theory) we need to have a prior awareness of the distinguisher or the cloak or the guise.

Let us formulate the first principle as follows:

> P1: If something x is presented in my awareness or knowledge, it is presented there under the cloak of a purported property.

Here 'the cloak of a purported property' is called a mode (prakāra). In what follows, I shall call the purported property the qualifier, and the object x the qualificand. This

B. K. Matilal and J. L. Shaw (eds.), Analytical Philosophy in Comparative Perspective, 373–391.

qualifier-qualificand structure is regarded as implicit in the content of any cognitive episode that has a claim to be regarded as knowledge (in the episodic sense). An exception can be made for the cases of 'simple' awareness or the so-called non-constructive (nir-vikalpa) awareness (if there are any) where the object shines in its nakedness, in its guiseless glory. But such awareness is not called knowledge in Nyāya.

The above principle, if true, has several implications. First, if seeing or perceiving is also a species of knowing, then we have to say that all seeing is seeing-as. It has been generally held by the Buddhist epistemologists that there are pure cases of seeing where we see the object in its nakedness, the unique particular datum untinged with any qualifier. But such a case of seeing or sensing would not be counted, under the view we are examining, as a 'distinguishing' cognitive episode, which, when it is proven to be correct, may be claimed to be a knowing episode.

Secondly, the above principle seems to suggest that the usual distinction which philosophers make between our knowledge-that and knowledge-what (knowledge of things) does not present two irreducible alternatives. For, knowledge-that can be taken to be more fundamental, and in all cases of so-called knowledge-what there is, it may be argued, an implicit propositional structure: *F*(*a*) or *a* as qualified by or distinguished by *F*-ness. We may say that someone's knowledge of *F* is actually unpackable as his knowledge of what is an *F*, i.e. that which is distinguished by the purported property *F*-ness. (In this context, I am ignoring presumably the complication due to the fact that in reprsenting a knowledge-that in language, we invariably introduce an 'assertion' element in it. For the moment, we focus upon only the content of such 'assertions'.)

In view of the point just made, my knowledge of the pen I am now writing with is to be spelled out, in the minimal sense, as my knowledge of the thing *as* a pen. Alternatively, it could be my knowledge of the thing as a solid substance, as a shiny object, as an object made of silver and so on. But there cannot be any knowledge of the thing as totally unqualified (even 'thing-hood' would be a qualifing character in this view, and hence I can know it simply *as* a thing).

I may verbalise my knowledge as "a pen", but this would only be an elliptical expression in this theory: "(It is) a pen" or "Something characterised by pen-hood". Here is an indication of how a linguistic expression, in this theory, would yield its *meaning*. The word "pen", even when it is in a referential position in an expression, would not present the object directly--it would present the particular as qualified by the purported property, penhood.

A pragmatic transformation of the above principle (P1) in the context of a word generated awareness would be

> P'2: If something x is denoted by a word, then in the awareness (of the object x) generated by the word (in the hearer) the object x is presented under the cloak of a purported property.

This is acceptable provided we add the following explanatory notes. (a) We may change "denoted by" to "designated by" if the word is a singular term or a proper name. (b) 'A word denotes x' means that at the utterance of the word an awareness of the object x is generated in a competent hearer by the familiar causal law. (c) 'The familiar causal law' refers either to the fact that the word has been learnt in a "dubbing" situation and is thereby infused with a 'power' (śakti) to denote an object or that the word is learnt in some other way as infused with the 'power' to denote the same way. And by 'word-generated awareness' we mean that the word being uttered generates in the hearer the relevant awareness of the object by virtue of the said 'power'. The upshot of all this is that what is denoted by the word is presented in the awareness of the hearer under a cloak or a mode. I shall come back to talk about the word-generated awareness in section 4. For the moment let me concentrate upon awareness in general and perceptual awareness in particular.

There is a well-entrenched philosophical view that says that perception of the thing per se, of the pure, 'uncoloured' existent, is what yields a piece of knowledge (pramā) in the basic sense, for it has not only psychological certainty, but also logical incorrigibility. If the (above) second point is conceded, it would go directly against such a contention of the epistemologists. Diṅnāga and many others have claimed that a sensation of blue, for example, when the object shines in its own glory and the subject grasps it directly, the purity of the object being not contaminated by anything, yields a very respectable piece of knowledge. If such an awareness cannot be called knowledge, what else would there be left for us to rely upon as the foundation of our knowledge?

In reply, the upholder of the above position may say that he does not dispute the fact that such episodes of pure sensation arise in us. What is in dispute is the cognitive character of such episodes. They are not pieces of empirical knowledge. To call them knowledge would be an inadvertent smuggling of an a priori character into what are essentially episodes of empirical knowledge. It is an important mark of our empirical understanding that it is liable to error. If we preclude any chances of our being in error with respect to the sensing of an object (and this is what 'logical

incorrigibility' apparently means), then we forfeit its claim to be considered as a piece of empirical knowledge. The general idea behind this point is this: In order to be either true or false, an awareness must be an awareness of something _x_ as _F_, i.e. as distinguished by a qualifier, _F_-ness, or (simply) by _f_. A true awareness is knowledge, and a false one an error, according to the terminology of the theory we are discussing here. Notice that in this theory, if I say "I sense _F_-ly" I may be reporting a cognitive event where the object _grasped_ is a particular sensation _x_ which is distinguished by the distinguisher, _f_. In the same vein, in most cases of my awareness of blue, it can be unpacked as that of the wall as blue, the sky as blue, or the _śārī_ of my beloved as blue. Only with such unpackings, the relevant episodes can make knowledge-claims. For they are, in each case, liable to be false. For example, if the wall is white, the sky is colourless and my beloved wears a red _śārī_, I would be in error in each case. Such knowledge claims can be entertained only when the facts are liable to be different.[1]

We normally do talk about knowledge of things, and hence such usages must be explained. According to this theory, when I know the thing in my hand as a dot-pen, I can talk about this knowledge-episode in either of two ways: I can call it either my knowledge of the dot-pen or my knowledge of the purported property, viz., being a dot-pen; and in either case we may pick out the same episode. My knowledge of the thing thus seldom arises in isolation _from_ my knowledge of the _way_ the thing is presented in my awareness.

P1 raises the following problem. If I know _x_ as distinguished by a property, _f_, then _f_ is also a part of what I know. This implies that _f_ also _floats_ in my knowledge as much as _x_ does, and hence one can argue that we need a further distinguisher for qualifying _f_. If I knew a piece as a piece of gold, then _being gold_ is also what I know. It would be absurd to claim that I do not know what GOLD is or what _being gold_ is like, and yet I know this piece to be gold. If this claim is right, then by our P1 we must say that if I know _being gold_ (or gold-ness), I should know it under a further characterisation. This leads to the peril of infinite regress: If _x_ figures in my awareness by way of being gold, and being gold figures there by way of being something else, then there will be no stopping. To avoid this problem, an exception to P1 is formulated.

> E1: When I am aware of an ultimate universal, a _simple_ property (a _jāti_ or an _akhanda upādhi_), I may be aware of it as such (unqualified). But if in the _verbal report_ of the awareness the simple property is mentioned by a name it needs to be

presented there under a further cloak of a purported property.[2]

Two observations are needed in this context. First, our sense of such expressions as "purported property", "ultimate universal" and "simple property" may be taken to be ontologically neutral. Such properties may or may not be separately real, or existent in the mind-independent objective world. It may be that there are only chairs in this world, but no separate thing called 'chairhood'. Our talk of chairhood is restricted only to its being a recognizable distinguisher (<u>viśeṣaṇa</u>). It is significant that Nyāya does not make any distinction <u>in this context</u> between a real (objective) universal (<u>jāti</u>) and a nominal universal in so far as they play the logical role of 'simple' properties. A simple property is ultimate in the sense of being a property that is further unanalysable (unbreakable).

This leads to the second observation. We can have, according to E1, a very direct "communion" with such simple properties, an uncoloured, non-mediated acquaintance. The distinction between such a knowledge of a simple property and the pure sensation of the uncoloured, unpropertied, naked object (about which knowledge-claim has been previously denied) is this: The former is being called here 'knowledge' only in virtue of its being an integral part of a knowledge-episode, such as knowledge of <u>x</u> as distinguished by a qualifier <u>f</u>, a simple property; but the latter has to stand apart, and be counted. Whether or not a purported awareness of a simple property can also stand apart and be counted (i.e. be a separate episode) is a controversial issue to which I shall now turn.

2. AWARENESS OF THE QUALIFIER

Our second principle mentioned in the beginning says that if there arises a qualificative awareness where an object <u>x</u> is distinguished by a qualifier <u>f</u>, in the above manner, one of its necessary pre-conditions is a prior awareness of <u>f</u>, the qualifier.

P2: To generate an awareness in which the object <u>x</u> is presented as qualified by <u>f</u>, a prior awareness of <u>f</u> is needed.

In order for me to be able to characterize or qualify <u>x</u> by <u>f</u> in my awareness, or attribute <u>f</u> to <u>x</u>, I must be in possession of an awareness of <u>f</u> prior to it. Unless I know what 'blue' or <u>being blue</u> is, I cannot judge something to be blue.

This principle is appealed to by the Navya-Naiyāyikas in order to settle an intricate controversy in their traditional

theory of perception and knowledge. Previously we have noted the Nyāya ambivalence about the status of a pure, pre-linguistic, conception-free, sensory grasp of the object in its theory of perception. Nyāya denies knowledge-hood of such episodes, and argues further that we are never 'consciously' aware that such a sensation has arisen. In other words, one is never aware that one is sensorily aware of anything in this pre-linguistic, conception-free manner. Powerful arguments of Bhartṛhari have well persuaded the Naiyāyikas to recognize the 'word-impregnated', conception-laden nature of our awarenesses, our thoughts, which are properly formed, employed and communicated.[3] But the Older Naiyāyikas did talk about a non-conceptual sensory grasp. In order to resolve the issue whether or not such graspings do arise in us, the New Naiyāyikas (Udayana, Gaṅgeśa etc.) emphasized the fact that only an inference can help us in deciding this matter. Roughly, the procedure is this: if we believe in P2, then my perceptual awareness of something x (say a piece of metal) as qualified by f (being gold, or goldness) must be preceded by my awareness of f (what it is to be gold, or goldness). Combining P1, E1 and P2, we may then say that as long as goldness is a 'simple' property according to our definition, it is possible for us to have a non-qualificative, non-mediated, perceptual awareness of goldness prior to the proper judgmental perception. The word 'perceptual' need not raise our eyebrows, for Nyāya maintains that if the individuals are perceptible, the so-called universals or simple properties residing therein may also be perceived, unless there is some stronger reason to believe them to be imperceptible. The only other difficulty in this is that universals like goldness or cowness are thought to be 'abstract' in some sense, while their locations, the individuals, are 'concrete' and hence, perceptible. But this 'concrete-abstract' division will cut no ice with Nyāya, for such a distinction does not exist in this system. It will be further argued that just as I can see that the chair has four legs (which is nothing but the property four-leggedness), I can also see that this has chairness. Hence it can be called perception of chairness mediated, of course, by conception. But enough of this digression. Let me return to an examination of P2.

We may say that we need simply a conception of f (goldness), in order to be able to have an awareness of x as qualified by f. This can be supported by arguments of the following kind. Unless some awareness (or conception) of fire is already present in a person, he cannot infer that the hill has fire there because there is smoke there. If I have never been aware of what it is for an object to be a camel, I cannot certainly be aware, all of a sudden, of an object as being a camel. Even a colour-blind person must understand the meaning of "green" as a compatible colour, in order to be able to

comprehend what is meant by the sentence "This is green". In this way, only a prior conception of the qualifier as such needs to be postulated for making a qualificative awareness possible.

It may be countered that the above argument is faulty. For, even if we concede the point about the conception of the qualifier in the case of a non-perceptual awareness, in perception such a prior stage may not be needed. For, it is argued, the contact of my sense-organ with the qualifier, 'colour', would be enough to generate the awareness of the object as coloured. Gaṅgeśa, in fact, has conceded all this, if only implicitly.[4] For him, all that we need is a notion of the qualifier f, somehow presented prior to our being aware of something as qualified or distinguished by f. In some cases, such a presentation may be made possible through the revival of some memory-impression. In the case of inferred knowledge or 'verbal' (sentence-generated) knowledge, such a requrirement is supplied by what is technically called sādhyaprasiddhi (literally, familiarity of the predicate property). I can neither infer something to be an abracadabra, nor understand the meaning of, i.e., have a knowledge from the utterance of, the sentence "It is an abracadabra", unless I am already familiar with what it is for a thing to be an abracadabra. But cases of perception are certainly different.

An example is considered by Gaṅgeśa to get around the difficulty in a perceptual situation. I shall however reformulate the example to make the point clear. Suppose, a disc has just turned blue, and I am looking at it. Further suppose that it has a particular blue-tint, the like of which I have never seen before. Now, for Nyāya, the qualifier can be either a universal property or a particular one; and in this case, the particular blue-tint would be the relevant qualifier. (But such a particular has to play the role of a property, i.e. it has to become universal-like, in the context a propositional combination: it has blue-tint b_1.) The argument continues by pointing out that I would in this case first see the blue-tint, the particular, in conception-free awareness before I could become perceptually aware that the disc is qualified by that particular blue-tint. In other words, I would have no other access to the idea of that blue-tint (obviously it cannot be remembered, because it has never been experienced before) except what my senses will yield. Here, therefore, we have a possible case for a simple perceptual awareness of a 'simple' entity, although this entity here is not a universal. The opponent may still argue that the particular blue-tint, in effect, will be seen, according to our P1, as a blue-tint, this means that this will be a complex awareness of the particular colour as being blue. I cannot simply have an awareness of the tint without seeing it as

blue. The notion of being blue, an ultimate universal, would in that case be supplied by a memory-revival. But we can still say that this memory-revival, if it has occurred, is occasioned in such cases (compare: prāthamika-go-pratyakṣa - Gaṅgeśa, jāgarādya-kālīna-go-pratyakṣa - Vardhamāna) by the senseory apparatus (sense-object intercourse); and since the object, the particular blue-tint, is only visually given, it would not be a case of remembering, but a case of simple perceptual awareness.[5]

A follower of Bhartṛhari may continue the debate in another way. It may be claimed that the notion of cowness, horseness or goldness may be congenital to us, and this will be postulated on the basis of the pan-Indian belief in transmigration and previous births. The notion of many 'simple' properties, the universals or natural kinds, may be only a retention of memory from previous births. This is the nearest equivalent of the 'innate idea' theory in the Indian context. Unlike the Western rationalists, the Indian thinkers never say that there are some innate ideas in us; instead, the hypothesis that is put forward in this regard is that the ideas which seem to be congenital (innate) are acquired through experiences over countless previous existences. Hence, when a child first recognizes a cow as a cow, he may be aided simply by the memory-traces of his previous births. Some verbal instruction is however needed to revive his memory-traces. This cannot ultimately rule out a primary perceptual experience at some time in the past. In any case, we can ignore the hypothesis of previous births in this context, and endorse the nearly conclusive argument of Gaṅgeśa that at some point there could be cases of pure perceptual grasp of the simple properties, the qualifiers as such, which will then precede some of our qualificative perceptual knowledge of the object as qualified by such properties. Such 'simple' properties are mostly universals or properties of natural kinds, but in some cases they may even be non-universals, which are then combined with another particular in a qualificative awareness. It is the awareness itself that combines two such particulars into a qualifier-qualified tie.

3. PREDICATE-PROPERTY AND THE QUALIFIER OF THE SUBJECT

It is important to distinguish the present issues from certain parallel problems of distinguishing between the immediate and the mediate perception, the direct and indirect knowledge--problems that are usually current in philosophical writings on perception. We are trying to outline here a general theory of cognition or awareness, following the Nyāya principles; and this, I think, may throw some light on the intricate problems connected with meaning and denotation of words.

A critique of P2 can be formulated in a very general way. If a prior awareness, in some form or other, of the qualifying entity is necessary for the arising of a cognitive awareness (a supposition, or a knowledge) of an object (the qualificand) as qualified by such a qualifier, i.e., as distinguished by a purported property, why is it not equally necessary to have a prior awareness of the object itself, the qualificand? The general principle of thought seems to be that it is not possible for a person to have a thought about something that it is _F_ or that it is qualified by _F_-ness, unless he knows which particular individual in the world he is thinking about. If, for example, I suppose or judge or know that a particular cow is white, then it is not only needed that I should already possess an awareness or knowledge of what it is for something to be white, but also that I should know that such a cow exists. Hence if it is emphasized that a prior awareness of white colour is needed, it may be equally emphasized that a prior awareness of such a cow is also needed.

This, however, is not a criticism of the Nyāya view, for, with a little twist, this criticism can be turned into a clarification of the Nyāya position. Let us consider the verbalized version of a knowledge-episode such as

"_A_ cow is dark."

(I shall not consider the more usual "_The_ cow is dark". For one thing, the Sanskrit philosophers seldom discuss such formulations. For another, this presupposes that the object _x_ is identified in more than one way: (i) by being qualified by cowness, as well as (ii) by being a previously identified object in the discussion or context.) Nyāya would say that the qualifier here is not only the dark shade or being dark, but also cowness. Cowness is called (in this context) the _dharmitāvacchedaka_, the delimiting character of the object _x_ to which another qualifier, or attribute, has been attributed. We have to know both qualifiers in order to know the object _x_ as an object qualified by them. Under the usual interpretation, the above mentioned knowledge would be explained as that of an object _x_ which is first qualified by cowness and then, being so qualified, it is further qualified by a dark shade, and this dark shade, in its turn, is qualified by being a dark shade (a universal, a simple property). Notice that being a dark shade is not a qualifier of _x_, rather it qualifies one of the qualifiers of _x_, and in this respect, that qualifier of _x_ is playing the role of a qualificand, another object, viz. the dark shade.

One of the implications of the above critique and reply is that P1 actually leads to a theory of identification of objects through descriptions or information about them. I cannot identify an object unless I already possess some

information about it. Part of this information may be perceptually given as in the case of being a cow when a child (à la Gaṅgeśa) first perceives a cow and then identifies it as a cow. Here "perceives a cow" would be interpreted by Naiyāyika as the direct grasp of the three-dimensional 'cow-substance' (cf. go-pinda) plus the cow-feature--the universal--but without any awareness of their connectedness. Similiarly "identifies it as a cow" is to be interpreted as knowing or supposing that it is a cow. But the first direct grasp of two discrete (unconnected) entities is too quick to be captured in the person's introspective awareness (anuvyavasāya). The direct grasp, if it arises at all, arises as almost an 'unconscious' awareness.

The (second) structured awareness is however not unconscious. Moreover its structure (a as F) presupposes a prior awareness of the qualifier, the attribute cowness, which is supplied by the first awareness here. It is however not emphasized in Nyāya that the said structure would need also a prior awareness of the subject entity, the 'cow-substance' (pinda). Just as the (unconnected) cowness is present there in the first awareness the awareness of the (again unconnected) cow-substance is also present there as an epistemic fact. But while the former is also required to contribute to formation of the structural content of the introspectable awareness (the second), the latter is not necessary in the same way. For although it is not possible for a person to have a thought about something that is F unless he knows which particular individual in the world he is thinking about, to have a (visual) perceptual awareness that something is F it is not needed that one must have a prior knowledge of the subject entity as such. The perceptual event itself would identify (single out) the subject entity and qualify it with the available qualifier at the same time.

We identify (pick out, single out) an object x in various ways, and a particular way (a particular mode of singling out) in a given context would be called the dharmitāvacchedaka, the delimiting character of the qualifiable object. These WAYS or modes are usually governed by bits and pieces of information the person has gathered about the object. In a perceptual context however the object may be singled out by the perceptual event itself without the ostensive aid of such bits and pieces of presumably previously gathered information in order that it may be instantly qualified by the available qualifier (cowness). Such a qualificative awareness would be called nirdharmitāvacchedaka jñāna, an awareness where the qualifiable object does not appear under any (other) distinguisher except the main qualifier, cowness. Roughly speaking, this awareness associates F with the naked object identified as such.

Most bits and pieces of information are perceptually gathered continuously by a person, and retained in his possession until some later time. Another important source of information is the exercise of reason upon the existing information. A third source is speech or language where the speaker wishes to transmit the information already in his possession. In this way all the pramāṇas can be seen as information-gathering instruments to help us identify the object. We would presently see how, according to Nyāya, a word would transmit information and help the hearer identify or know the object.

One further point before I proceed any further. When following Nyāya I am talking about an object that floats in our awareness, I am not talking about an idea in the mind or even what is called the 'content of consciousness'. Also I am not talking about a so-called mental image. For Nyāya, there is no such thing, no "veil of ideas" between us and the things outside, and no mental images. In other words, the object is not 'mental' unless we are talking about such objects as internal states or psychological events, e.g. a desire, a feeling of pain, etc. According to Nyāya, an external object, when it is connected with awareness, stays the same--no novel object, no image, not even an idea, is created in the awareness; when an awareness arises in the subject, we can only say that a connection, called cognizedness, has been established momentarily between the object and the awareness. Nothing more need be conceded here.

4. LANGUAGE-GENERATED AWARENESS AND MEANING

I shall now try to explain the nature of a language-generated awareness or what is called śābdabodha in Nyāya. It is the awareness presumably of the competent hearer from the utterance of words or sentences. Although we shall be concerned here only with words and what they signify for the hearer, these words are always to be regarded as parts of some sentence or other. The theory presupposes the existence of a language community where speakers utter words and sentences to convey thoughts, intentions, commands etc., and there are hearers who understand them from the knowledge they derive from such utterances. It is this knowledge that is called śābdabodha, and is distinguished from perception etc.

The process of acquiring this knowledge is described as follows. First, the utterance of words would directly generate knowledge of such words. This knowledge of the words leads to the awareness of the objects meant by them, and in this way, parallel to the grammatical connection of the words in the sentence, a knowledge of a connected meaning arises, which is called the product (phala) of the entire process. I shall concentrate here upon the knowledge generated by words

rather than sentences.

In order to have awareness of the objects meant from the knowledge of words, a special meaning-linkage between the word and its meaning is needed, and the hearer must be well aware of such a linkage in each case. We may call this meaning-linkage the 'denoting power of the word' (translating thereby the term *śakti* in Sanskrit). Besides the denoting power, the words may have another power by virtue of which it would help generating the knowledge of its meaning or the object meant. This is called 'metaphor' (cf. *lakṣaṇā*), which simply extends or contracts the denotative domain so that the right object may be picked up to suit the context or intention of the speaker. An example will make the point clear. The word "river" generally means a mass of flowing water. In our terminology, it presents to the hearer, through the denoting power, an object qualified by the property riverhood, i.e. being a mass of flowing water. But in the sentence "The village is on the river", the word has a metaphorical use. It presents to the hearer (if he understands idioms and metaphors of the language) an object that is not the river itself, but only the bank of the river. How? By virtue of its being connected with some river.

How does the word get its denoting power? Indian philosophers are divided on this issue. According to some, it is 'natural' to the word. According to Nyāya, however, the word is invested with the power either by God's decree (in the case of words where nobody knows how the meaning was fixed) or modern stipulative definitions. In either case, however, the structure of the stipulation is the same. It is of the form of a will, "From the word 'A' let us be informed of *B*", or "Let this word 'A' inform (us) about *B*". The competence of the hearer means that he has gathered information about such stipulations. Being infused with such stipulated denoting power, the word "cow", for example, will be competent to generate the awareness of the object cow, but if we follow our previous principle, P1, the object will float in our awareness only as being qualified by cowness, as being delimited by cowness.

To cut a long story short, the hearer must acquire the knowledge of such denoting power of words (cf. *śaktigraha*) either by watching (as a child does) the speakers and the hearers in the community, their actions, responses etc. (the child generally watches the behaviour of those giving commands, and those obeying or disobeying them, cf. *prayojaka-vṛddha* and *prayojya-vṛddha*) or through instructions from grammar, lexicon, intelligent guesses etc.

The knowledge of the denotative power, it is claimed, takes the following form "The word X is empowered to present...." What fills the gap here would be called *śakya*. For convenience, I will call it the object *meant*. There are

usually several views about what is exactly the object that is meant by such words as "cow". I shall refer to these three views. According to one, the word cow means the universal, 'cowness'. According to others, it is the individual cow, the object x, that is meant. According to Nyāya, the word "cow" means, i.e. presents in the awareness of the hearer in the way described above, what is qualified by cowness, the object x as distinguished or delimited by cowness. In this way, cowness functions as the delimitor or distinguisher of the object meant by "cow". The older grammarians used to call it the 'occasioning ground for the use of the word' (pravrttinimitta).

So far we have not said anythng about proper names and descriptive phrases. It is clear that in this theory the proper name cannot present the object meant to the awareness of the hearer in an unqualified manner. Jagadīśa has called such words pāribhāṣika.[6] In order to save P1, one can always think of some nominal property to qualify the object presented. In the case of descriptive phrases, the object is identified or presented to the awareness of the hearer as qualified by some purported property, some observable feature etc. called upalakṣana. For example, "cook" would mean an object, a person, which is marked off by the profession of cooking.

5. PROPER NAMES AND INDEXICALS

I shall conclude by noting several consequences of the theory of awareness and language mentioned above. First, it seems to me that there cannot be any genuine proper name in language, if the above theory is strictly adhered to. By a genuine proper name, I mean the name that will fulfil at least two conditions: (i) it names or identifies only one object, a single individual, (ii) it is directly latched on to the object such that the object would be presented as such, unqualified, to our awareness (as soon as the word is uttered) and not as qualified by some qualifier. In this regard, awareness due to the use of a proper name would resemble the purported sensory grasp of the 'naked' object, and we have already seen how difficult it is to make sense of such a non-qualificative awareness in the Nyāya view. Some older Naiyāyikas (as reported by Gadādhara) believe that a non-qualificative revival of the memory of the object per se is possible from the utterance of the name. But this would not be a śābdabodha or language-generated awareness, for there would be other elements in the actual language (e.g. use of a singular suffix etc.) so that the object would be presented in our resulting awareness as qualified by such singularity etc.

Another view is that all singular terms be introduced by definition (paribhāṣā, Jagadīśa), and the purported

defining property be the qualifying character when the object is presented in our awareness by such a term. But this will certainly lead to further problems, as Gadādhara points out, with regard to such definition-sentences, among other things.[7] For example, the definition-sentence

"The moon is the satellite of the earth"

cannot generate awareness in the usual way. For, the object, viz. the moon, is to be presented or identified per se so that the property, viz. being the satellite of the earth, can be attributed to it. But a sentence-generated awareness cannot be such that the qualificand entity figures in it as totally unqualified (nir-dharmitāvacchedakaka, Gadādhara). As I have already noted, a perceptual awareness may be different in this respect from a word-generated awareness.

A third suggestion is to let the uniqueness of the individual be the qualifying character when it is presented in our awareness by the utterance of the name. It is presumed that during the name-giving ceremony, the linkage is established between such uniqueness of the object named and the name itself. But, it is argued, such uniqueness in the final analysis is nothing but the individual itself (this is how I interpret Jagadīśa's comment: tad-vyaktitvañca tādātmyena saiva vyaktiḥ).[8] Hence, the suggestion does not get us any further.

Gadādhara's suggestion is that we have to treat proper names differently from general names in that the latter designate things as qualified by the purported properties, but the former designate the unqualified objects, although when such objects are presented in our language-generated awareness they may be distinguished by some variable property or other. We may identify Rāma sometimes as the eldest son of Daśaratha, sometimes as the husband of Sītā and sometimes as the king of Ayodhyā. The proper name is thus not directly presenting the object in the hearer's awareness. The important thing to remember is that such distinguishers are no part of the meaning or the denoting power of such names. In the above example, when I hear the sentence, the moon may float in my awareness as the bright thing I see in the evening sky, but this qualifier or distinguisher, although it serves to distinguish the object presented, is no part of the denoting power of the expression "the moon". The expression may directly designate the moon, but the hearer captures it in his awareness always under some variable guise.

There are two other related consequences. One is that, according to this theory, the language-generated knowledge must be informative. It is therefore claimed that tautologous sentences, such as "A is A" do not give rise to any such knowledge, for they do not contain any information. Such

sentences would then be non-significant (meaningless?) in this theory, because they do not impart knowledge. One of the criteria for the significance of a sentence (such that it generates a śābdabodha) is that the delimiting character (qualifying property) of the subject-entity should not be identical with the main qualifier, the predicate-property. "A is A" is a sentence where A-ness is both the predicate property and the character delimiting the subject. Hence its utterance does not generate knowledge in the hearer. This may be quite in keeping with the sentiment expressed by some philosophers that to say of any one thing that it is identical with itself is to say nothing.[9] But the Nyāya criterion is even stricter than what appears here. Of course, one may say that "statement of identity consisting of unlike singular terms may be non-idle and hence significant".[10] But Nyāya claims that this would not do. For, in the definition-sentence (see the above sentence) we have two unlike singular terms, but still its interpretation, according to Nyāya, runs into problems. Apparently the definition is supposed to supply the essential attribute of the moon and thereby give the delimiting character of the moon. But the subject-entity, the moon, is presented as qualified by the delimiting character, viz. being the satellite of the earth; and since it is the same property that we ascribe to it by the predicate expression, we face the same impasse as we do while dealing with "A is A". To resolve the impasse, Gadādhara introduced the important distinction between two types of distinguishers or properties used to identify or single out an object.

This brings me to the last consequence that I shall discuss here. A distinguisher or qualifier may identify an object in a totally irrelevnt way, or it may be very pertinent to the identification of the object in the context of the further attribution of properties to it. I may grab an object by hand (cf. śṛṅgagrāhikā vyavasthā) and say that it is F, or understand it as being qualified by F-ness; but the manner of identification, viz. grabbing by hand, is totally irrelevant to the object's being F. I can identify an object as the man with the whisky-glass in hand, and go on to say that he is a logician, where the distinguisher, viz. having a whisky-glass in hand, is totally irrelevant to his being a logician (cf. kākavad gṛham, jaṭābhiḥ tāpasaḥ). But if I go on to say that he is going to get drunk very soon, then at least the audience may understand (or misunderstand) that the identifying qualifier, having a whisky-glass in hand, is relevant. Nyāya underlines the distinction by calling the relevant distinguishers viśeṣaṇa, and the irrelevant ones, upalakṣaṇa. One can, in this way, identify an object by an irrelevant qualifier, and assign a proper name to it. This would mean that there may be an irrelevant qualifier of the object presented to the awareness of a hearer by the utterance

of a proper name. The connection between the name and the qualifier of the object presented by the name to the hearer is entirely arbitrary.

This expedient can also be used to explain part of the pu ement about two different uses of the same sentence such as "Smith's murderer is insane". On one occasion it is said when no murderer has been apprehended, and there is no suspect, the observation being made by looking at the way the murder has been committed. In the second, it is said when a person, say John, is already apprehended, and the observation is made on the basis of his behaviour. In both cases, we have the same qualifier to capture or identify the subject-entity, the object; but in the latter case, this property, that of being seemingly the murderer of Smith, is an irrelevant distinguisher or qualifier, a *upalakṣana*, while in the former, it is a relevant one, a *viśeṣana*.

How does the distinction between *viśeṣana* and *upalakṣana* compare with the traditional Western distinction between essence and accident? A brief comment on this point is in order here. Both *upalakṣana* and *viśeṣana* are called 'distinguisher' (*vyavartaka*), and this undoubtedly underlines the epistemological character of the pair. In this context Nyāya was not interested in the ontological question of essence and accident. On the other hand the Naiyāyikas in general work within the framework of a metaphysic of substnace, in which context, however, the notions of essence and accident are pertinent. Their doctrine of objective universal seems to be of a piece with the notion of essence.

A *viśeṣana* or what I have called a *relevant* distinguisher need not however be an essential attribute in any ontological sense but simply be an attribute required for the application of a term. It is strictly related to our language, our vocabulary. It is meaningless to speak of a *viśeṣana* or relevant distinguisher except relative to the object's being denoted by a particular term. Sometimes it might coincide with the ontological essences or the objective universal but that is not the issue here. An *upalakṣana* or what I have called an *irrelevant* distinguisher is not even related to the general terms (or singular terms) commonly used to denote (or designate) such objects. For example, the distinguisher that distinguishes the object designated by "the moon" in the above sentence is not related to the particular term, the designator itself. It is only an ad hoc or ad hominem way of distinguishing the object by any suitable property.

Gadādhara has used this expedient quite convincingly to resolve some of our pu es about pronouns, or indexicals like "it" ("*tat*"), to which I can make only a very brief reference here. One of the problems with the use of "it", as Nyāya sees it, is that it presents or captures different objects at each time we use the expression, and we cannot talk about a

property "it-ness" (in the way we talk about cow-ness for different cases of using the word "cow") as the qualifier by virtue of which the word "it" would present the object each time to the awareness of the hearer. On the other hand, "it" cannot be regarded as one of the homophonic words (accidents of most natural languages) with different meaning in each time of its use. It is also not a proper name where the speaker could possibly stipulate as to what it means. There is some intuitive reason to believe that it is one of the 'natural' words of the language, like the word "cow" or "beast", where we need a common feature (such as cowness for "cow" and having a hairy tail for "beast") to capture each time the object it refers to by qualifying the presented object by such a qualifier.

One suggestion is that we take the property, viz. being in the thought of the speaker, as the purported common feature which will qualify whichever object the word "it" refers to in any of its uses. Thus, when a chair is meant by the use of "it", the object will be presented not as qualified by chairhood, but by being in the thought of the speaker. Similarly when a horse is meant, the object will be qualified by the same common feature. But this seems to be counter-intuitive. For, there are strong reasons for believing that in the following sentence the word "it" presents the chair, the object, as having a necessary connection with chairhood:

"There is a chair, bring IT."

For nobody would claim here that the hearer who understands the sentence could bring a non-chair, something else that may be qualified by being in the thought of the speaker, and still be said to behave correctly. For successful behavior the hearer must be aware, from the utterance of "it", of the actual chair qualified by chairhood.

Gadādhara says that of course the "it" in the above sentence presents the chair, i.e. the object qualified by chairhood, in the awareness of the hearer, but the hearer's route to capture this object is rather an irrelevant qualifier or distinguisher, that of being in the thought of the speaker. The latter property is an irrelevant distinguisher or identifier of the object (an upalaksana) in the sense already described; after the object has been identified, its function as a distinguisher comes to an end. The said property certainly provides the unifying character needed for our acquiring the knowledge of the denoting power of "it", but it is not part of its meaning. It is only a catalyst-agent, being not relevant to what is said or what we may continue to say. After helping the hearer to identify the object of bringing, it ceases to be operative. There are, of course, many other problems that arise in the case of such indexicals, and Gadādhara discusses them to some extent.[11] But I forbear to enter into them here.

NOTES

1. This is an appeal to the usual Sanskrit dictum: sambhave vyabhicāre vā syād viśeṣaṇam arthavat (an adjective is significant provided it is possible for it to be true or false of the object to which it is applied).

2. This rider is needed for the Naiyāyikas say that a 'simple' property may be cognised as such provided it is not mentioned by name (ullikhyamāna) in the verbal report of the cognition. If the verbal report of an awareness mentions "cow" then we have to infer that cowness, which appears as a qualifier of the object mentioned by "cow", itself appears as such (unqualified) in that cognition. If however the verbal report mentions cowness by name we have to say that it has appeared in the relevant cognition under another mode or qualifier.

3. For Bhartṛhari's view on this point, see Vākyapadīya, Kāṇḍa I, verses 115-116 (ed. S, Iyer, Poona, Deccan College, 1966).

4. Gaṅgeśa, Tattvacintāmaṇi, Pratyakṣa khaṇḍa (ed. Pt. K. Tarkavagisa, Calcutta, Bibliotheca Indica, 1897), Nirvikalpa-vāda, p. 817 ff.

5. Consult Gaṅgeśa, ibid., p. 817. For Vardhamāna, see his commentary on Udayana's Tātparya-pariśuddhi (Bibliotheca Indica, Calcutta, L.S. Dravid and V.P. Dvivedin eds., 1911), pp. 533-535.

6. Jagadīśa, Śabdaśaktiprakāśikā (Kashi Sanskrit Series 109, ed. Pt. D. Sastri, Beneres, 1973), p. 122 ff., verse 23.

7. For Gadādhara's view see his Śaktivāda (Kashi Sanskrit Series, Chowkhamba, ed. G.D. Sastri, Beneres) p. 54 f.

8. This is from Jagadīśa's comment on the Pakṣatā section of Tattvacintāmaṇi (see Gaṅgeśa, Anumāna khaṇḍa, ed. S. Nyāyopadhyaya, Chowkhamba, Beneres, 1906-1908).

9. L. Wittgenstein, Tractus Logico-Philosophicus, London, 1963, 5.5305.

10. W.V. Quine, Word and Object, Cambridge, Mass., 1960, p. 117.

11. While discussing the problems connected with such singu-

lar terms as "ākāśa" (= the sky), Gadādhara discussed the problem of meaning connected with such pronouns as "it", "that", "we", "I" and "you". A discussion of these issues will take us far beyond the scope of this paper.

INDEX